

TRIPHOSPHATE OLIGONUCLEOTIDE MODIFICATION REAGENTS AND USES THEREOF

RELATED APPLICATIONS

This application is related to co-owned U.S. utility application, entitled "FUNCTIONAL BIOPOLYMER MODIFICATION REAGENTS AND USES THEREOF", to Schwartz et al., filed August 1, 2000 (Attorney
5 Docket No. 37154-751), and to U.S. provisional application Serial No. 60/191,186, entitled "HYDRAZINE-BASED AND CARBONYL-BASED BIFUNCTIONAL CROSSLINKING AGENTS", to Schwartz, filed March 22, 2000. The disclosures of the above-referenced applications are incorporated herein by reference in their entirety.

10 FIELD OF THE INVENTION

Monomers and methods for preparation, detection and immobilization of macromolecules, including biopolymers, and for preparation and detection of macromolecular conjugates are provided. The monomers for use in the methods provided herein include hydrazino,
15 oxyamino and carbonyl substituted nucleoside triphosphates. Biopolymers, including oligonucleotides, possessing hydrazino, oxyamino or carbonyl modifications are also provided.

BACKGROUND OF THE INVENTION

Polymerase chain reaction (PCR) expression of oligonucleotides is a
20 powerful tool in molecular biology. The PCR product is generally labelled to aid in detection. Various methods of direct and indirect labelling of the PCR product have been developed. Direct labelling involves the incorporation of a labelled monomer during enzymatic synthesis, while indirect labelling refers to post-synthetic introduction of the label.

25 For example, methods for direct labelling of PCR products are known (see, U.S. Patent Nos 5,242,756 and International Patent Application Publication No. WO 99/65993) In such methods, modified nucleoside triphosphates are incorporated into oligonucleotides during amplification; and conjugation of a xanthene or cyanine label to a

nucleoside triphosphate is effected via an amide bond formed with the modified nucleoside triphosphates. Other methods involve conjugation of a biotin or metal chelating label to an oligonucleotide via an amide or hydrazide bond is used to form a desired oligonucleotide-label conjugates (see, U.S. Patent Nos. 4,707,440 and 4,889,798). Another method for indirect labelling of oligonucleotides involves incorporation of a boronic acid containing nucleoside triphosphate into an oligonucleotide during enzymatic synthesis. Complexation of this boronic acid modified oligonucleotide with a hydroxamic acid derivatized label provides the desired labelled oligonucleotide (see, U.S. Patent No. 5,876,938).

These methods of oligonucleotide labelling, however, are limited by the lack of stability of the label to the amplification reaction conditions, the inability to selectively and specifically incorporate multiple labels, and the instability of succinimidyl esters. Preparation of activated functionalities, such as succinimidyl esters or maleimides, of labels can be costly, and not even possible in some instances, particularly under PCR conditions.

Thus, due to the limitations of currently available methods as described above, there is a need for efficient methods for labelling of oligonucleotides. Therefore, it is an object herein to provide monomers and methods for labelling of oligonucleotides without the need for post-synthetic modification of the oligonucleotide. It is also an object herein to provide monomers and methods for enzymatic synthesis of modified biopolymers, including oligonucleotides, that can be specifically labelled. A further object herein is to provide the resulting modified oligonucleotides.

SUMMARY OF THE INVENTION

Oligonucleotide monomers containing hydrazino, oxyamino, or carbonyl groups that can be incorporated into an oligonucleotide chain during enzymatic oligonucleotide synthesis are provided. Methods for immobilization and conjugation of biopolymer first components,

particularly oligonucleotides, containing hydrazino, oxyamino, or carbonyl modifications are provided. The monomers are triphosphate nucleoside derivatives that can be incorporated into an oligonucleotide first component during enzymatic synthesis, including, but not limited to, synthesis by polymerase chain reactions (PCR), reverse transcriptases, including, but not limited to, AMV reverse transcriptase, MMLV reverse transcriptase and superscript reverse transcriptase, and polymerases, including, but not limited to, Taq polymerase, DNA polymerase, Klenow fragment and T4 DNA polymerase. The resulting first components can then be used for any purpose for which oligonucleotides are used. They are particularly suitable for conjugation to a second component or immobilization on a surface. The monomers provided herein advantageously are readily incorporated into oligonucleotide chains, hence can be used in any application that involves or uses a nucleoside triphosphate, such as DNA and RNA sequencing, detecting, labeling and amplification methodologies. Monophosphate and diphosphate forms of the monomers as well as nucleic acid chains containing the incorporated monomers are provided.

The monomers provided herein are also useful as mass modifiers in DNA sequencing by mass spectrometry (see, e.g., U.S. Patent Nos. 6,074,823 and 5,547,835). The monomers can be incorporated into oligonucleotides for the accurate determination of base composition (Muddiman et al. (1997) Anal. Chem. 69:1543), and for the scoring of single nucleotide polymorphisms (SNPs) (Chen et al. (1999) Anal. Chem. 71:3118). The monomers can also be used to study the mechanisms by which ribozymes effect catalytic cleavage (Earnshaw et al. (2000) Biochemistry 39:6410). The monomers can be incorporated into antisense oligonucleotides to increase their resistance to enzymatic degradation (Verheijen et al. (2000) Bioorg. Med. Chem. Lett. 10:801), their overall potency (Flanagan et al. (1999) Proc. Natl. Acad. Sci. USA

96:3513) and the stability of their hybrids with the complementary RNA sequences (Compagno et al. (1999) J. Biol. Chem. 274:8191).

The monomers possess a triphosphate-ribose-nucleobase motif for recognition by the enzymatic catalyst.

5 Riboses for use in the monomers and methods herein are well known to those of skill in the art, and include, but are not limited to, fully hydroxylated sugars such as ribose, deoxyriboses such as 2-deoxyriboses, and dideoxy riboses such as 2,3-dideoxyribose.

10 Nucleobases for use in the monomers and methods herein are also well known to those of skill in the art, and include, but are not limited to, cytosines, uracils, adenines, guanines and thymines, and analogs thereof, including deaza analogs.

15 The monomers also possess, in addition to the triphosphate group, a protected or unprotected hydrazino, protected or unprotected oxyamino (-O-NH₂), or carbonyl moiety for formation of a hydrazone or oxime linkage with an appropriately modified surface or second component. The hydrazino moiety can be an aliphatic, aromatic or heteroaromatic hydrazine, semicarbazide, carbazide, hydrazide, thiosemicarbazide, thiocarbazide, carbonic acid dihydrazine or hydrazine carboxylate (see, 20 FIGURE 1). The protecting groups are salts of the hydrazino or oxyamino group, including but not limited to, mineral acids salts, such as but not limited to hydrochlorides and sulfates, and salts of organic acids, such as but not limited to acetates, lactates, malates, tartrates, citrates, ascorbates, succinates, butyrates, valerates and fumarates, or any amino 25 or hydrazino protecting group known to those of skill in the art (see, e.g., Greene et al. (1999) Protective Groups in Organic Synthesis (3rd Ed.) (J. Wiley Sons, Inc.)). The carbonyl moiety can be any carbonyl containing group capable of forming a hydrazone or oxime linkage with one or more of the above hydrazino or oxyamino moieties. Preferred carbonyl 30 moieties include aldehydes and ketones.

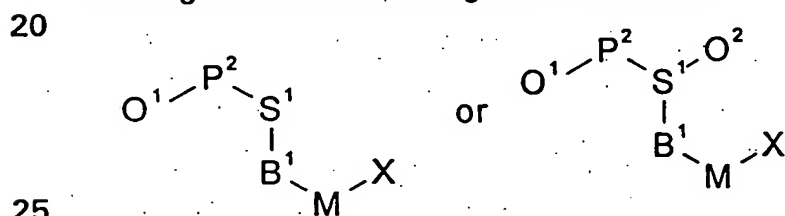
Second components include, but are not limited to, macromolecules, biopolymers as defined herein, polymers including, but not limited to, polyamines, polyamides, polyethers and polyethylene glycols, and other compounds of interest herein for use in assays, kits, diagnostic arrays, and the like, including, but not limited to, intercalators, vitamins, reporter molecules, cholesterol, phospholipids, biotin, phenazine, phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, dyes, antibodies, haptens, antigens, enzymes, and detection reagents including, but not limited to, fluorophores, metals including, but not limited to, gold, metal chelates, chromophores, fluorophore precursors and chromophore precursors, that possess or are modified to possess a hydrazino, oxyamino or carbonyl group that is complementary to the carbonyl, oxyamino or hydrazino group of the oligonucleotide of formula (II) for formation of hydrazone or oxime linkage. Dendrimeric compounds that possess a plurality of detectable groups, including, but not limited to, reporter molecules, fluorophores, chromophores, fluorophore precursors and chromophore precursors, are also contemplated herein as second components. Fluorophore precursors and chromophore precursors are compounds that react with the hydrazino, oxyamino or carbonyl group of the modified oligonucleotide to form a fluorogenic or chromogenic group for analysis. Such groups are preferred due to the absence of background noise in the resulting assay. Preferred conjugates include those containing a hydrazone linkage.

In one embodiment, the monomers for use in the methods provided herein are capture nucleoside triphosphates (cNTPs) that have formula (I): $P^1-S^1-B^1-M-X$ or a derivative thereof, as defined herein, where P^1 is a triphosphate group, as defined herein; S^1 is a ribose, a deoxyribose or a dideoxyribose; B^1 is a nucleobase; X is a protected or unprotected hydrazino group, a protected or unprotected oxyamino group, or a carbonyl derivative, where the protecting group is a salt or any amino or hydrazino protecting group

- known to those of skill in the art; and M is a divalent group having any combination of the following groups, which can be combined in any order: arylene, heteroarylene, cycloalkylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^1)$, $P(D)_b(ER^1)$, $N(R^1)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^1 ; D is S or O; and E is S, O or NR^1 ; each R^1 is a monovalent group independently selected from hydrogen and M^1-R^4 ; each M^1 is a divalent group independently having any combination of the following groups, which groups can be combined in any order: a direct link, arylene, heteroarylene, cycloalkylene, $C(R^5)_2$, $-C(R^5)=C(R^5)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^5)$, $P(D)_b(ER^5)$, $N(R^5)$, $N(COR^5)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^5 ; D is S or O; and E is S, O or NR^5 ; R^4 and R^5 are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, $SiR^6R^7R^8$, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; R^9 and R^{10} are each independently selected from hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heteroaryl, heteroaralkyl and heterocyclyl; R^2 and R^3 are selected from (i) or (ii) as follows: (i) R^2 and R^3 are independently selected from among hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl and heteroaryl; or (ii) R^2 and R^3 together form alkylene, alkenylene or cycloalkylene; R^6 , R^7 and R^8 are each independently a monovalent group selected from hydrogen, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; and R^1 , R^2 , R^3 , R^4 , R^5 , R^6 , R^7 , R^8 , R^9 and R^{10} are optionally substituted with one or more substituents each independently selected from Z,

wherein Z is selected from alkyl, alkenyl, alkynyl, aryl, cycloalkyl, cycloalkenyl, hydroxy, $S(O)_hR^{20}$, $NR^{20}R^{21}$, $COOR^{20}$, COR^{20} , $CONR^{20}R^{21}$, $OC(O)NR^{20}R^{21}$, $N(R^{20})C(O)R^{21}$, alkoxy, aryloxy, heteroaryl, heterocyclyl, heteroaryloxy, heterocyclyloxy, aralkyl, aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, aralkoxy, heteroaralkoxy, alkoxy, carbamoyl, thiocarbamoyl, alkoxy, carboxyaryl, halo, pseudohalo, haloalkyl and carboxamido; h is 0, 1 or 2; and R^{20} and R^{21} are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, trialkylsilyl, dialkylarylsilyl, alkylidiarylsilyl, triarylsilyl, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy, amino, amido, alkylamino, dialkylamino, alkylarylamino, diarylamino and arylamino.

Macromolecules, particularly biopolymers, including oligonucleotides, that are modified by incorporation of the above monomers are provided. Thus, in certain embodiments, provided herein are oligonucleotide analogs of formula (II):



or a derivative thereof, where O^1 and O^2 are each independently oligonucleotides or analogs thereof, such as protein nucleic acids (PNAs); P^2 is a phosphodiester group, resulting from the coupling of a compound of formula (I) with O^1 or analog thereof; and S^1 , B^1 , M and X are selected as above.

Such oligonucleotide analogs are useful as antisense drugs, cis-elements acting as regulators of gene expression, and substrates for RNA

binding proteins such as HIV-1 Rev and Tat (see, e.g., Agrawal (1999) Biochim. Biophys. Acta 1489:53; Crooke (1999) Biochim. Biophys. Acta 1489:31; Gryaznov (1999) Biochim. Biophys. Acta 1489:131; and Morishita (2000) Nipp. Yakuri. Zass. 115:123).

- 5 Oligonucleotides covalently immobilized on a solid surface, as described herein, are also provided. The oligonucleotides are of formula (II), and are immobilized through a covalent hydrazone or oxime linkage. The solid surface is modified to possess a hydrazino, oxyamino, or carbonyl group that is complementary to the hydrazino, oxyamino or
 - 10 carbonyl moiety of the oligonucleotide of formula (II) for formation of a hydrazone or oxime linkage. In preferred embodiments, the oligonucleotides are immobilized through a covalent hydrazone linkage.
- The immobilized oligonucleotides, as well as non-immobilized forms, are useful for a variety of purposes known to those of skill in the
- 15 art, including, but not limited to, preparation of microarrays, diagnostic probe assays, DNA or RNA amplification, for example by solid phase polymerase chain reactions (PCR), molecular computing (see, e.g., Adleman (1994) Science 266:1021-1024; Kari (1997) Mathematical Intelligencer 19:9-22; Frutos et al. (1997) Nucleic Acids Res. 25:4748;
 - 20 Smith et al. (1998) J. Comp. Biol. 5:255; Liu et al. (1998) J. Comp. Biol. 5:267; Frutos et al. (1998) J. Am. Chem. Soc. 120:10277; Wang et al. (1999) Biosystems 52:189-191; Liu et al. (1999) Biosystems 52:25-33; Liu et al. (2000) Nature 403:175-179; European Patent Application Publication No. EP 0 772 135; Reed et al. (June 2000) Scientific
 - 25 American:86-93), molecular addressing (Niemeyer et al. (1994) Nucl. Acids Res. 22(25):5530-5539), DNA and RNA sequencing methods, including mass spectrometry methods (see, e.g., U.S. Patent Nos. 6,074,823 and 5,547,835), nucleic acid diagnostics, including SNP and other polymorphism analyses and detection methods, and in studying the
 - 30 molecular electronics of DNA (see, e.g., U.S. Patent Nos. 6,071,699, 6,066,448, 5,952,172 and 5,824,473).

The oligonucleotides can also be used in PCR-based sequencing methods, particularly solid phase sequencing methods using the immobilized oligonucleotides (Mustajoki *et al.* (1997) Genome Res. 7:1054). They can also be used to measure the interaction forces
 5 between single strands of DNA (Lee *et al.* (1994) Science 266:771), in solid phase-mediated transfection of oligonucleotides (Bielinska *et al.* (2000) J. Biomaterials 21:877), and in solid phase cloning to create libraries suitable for direct solid phase sequencing (Hultman *et al.* (1994) J. Biotechnol. 35:229). The immobilized oligonucleotides can also be
 10 used in DNA chip technology to create arrays of oligonucleotides which are used to compare the qualitative and quantitative characteristics of gene expression profiles, mutations, insertions and deletions in normal and diseased states (De Benedetti *et al.* (2000) Int. J. Biol. Markers 15:1). They can also be used to identify and characterize the DNA
 15 binding site of DNA binding proteins (Roth *et al.* (1995) EMBO J. 14:2106; Carlsson *et al.* (1995) Anal. Biochem. 232:172). Immobilized oligonucleotides bound to Sephacryl S-500 particles via a CNBr-activation procedure can be used to assemble extended DNA duplexes by phosphorylation, ligation and restriction enzyme digestion of assemblies
 20 of annealed oligonucleotides in solid phase (Hostomsky *et al.* (1987) Nucleic Acids Symp. Ser. 18:241). The immobilized oligonucleotides can also be used in PCR, RT-PCR (Kozwicz *et al.* (2000) Appl. Environ. Microbiol. 66:2711; Blomqvist *et al.* (1999) J. Clin. Microbiol. 37:2813), transcription (Marble *et al.* (1995) Biotechnol. Prog. 11:393; Fujita *et al.*
 25 (1993) Biotechniques 14:608), ligation reactions (Filippov *et al.* (1990) Bioorg. Khim. 16:1045), and in studying DNA repair mechanisms (Salles *et al.* (1999) Biochimie 81:53).

In embodiments where X of the modified oligonucleotide is a hydrazino or oxyamino group, the solid surface can be modified to
 30 possess an epoxide, α -bromocarbonyl, maleimide, maleic anhydride, isothiocyanate or isocyanate group. Such solid surfaces can be prepared

by methods provided herein or other methods well known to those of skill in the art." For example, reaction of pentafluorophenyl 4-isothiocyanatobenzoate with an amino solid surface results in formation of an isothiocyanato modified solid surface. Some of these surfaces are commercially available from, e.g., Pierce (Rockford, Illinois), SINTEF Applied Chemistry (Trondheim, Norway), Rapp Polymere GmbH (Tubingen, Germany), and Dyno Particles AS (Trondheim, Norway). Reaction of the hydrazino or oxyamino group of the modified oligonucleotide with the epoxide, α -bromocarbonyl, maleimide, maleic anhydride, isothiocyanate or isocyanate group of the solid surface results in covalent attachment of the oligonucleotide to the solid surface.

In certain embodiments, particularly where X is an oxyamino group, the immobilized oligonucleotides are selected such that the solid surface is not modified with an aldehyde or epoxide group.

Oligonucleotide conjugates are provided. The conjugates are prepared from the modified oligonucleotide first components of formula (II). The modified oligonucleotide of formula (II) is reacted with a complementary derivative of a second component to form a hydrazone or oxime covalent linkage. The second components include, but are not limited to, biopolymers as defined herein, polymers including, but not limited to, polyamines, polyamides, polyethers and polyethylene glycols, and other compounds of interest herein for use in assays, kits, diagnostic arrays, and the like, including, but not limited to, intercalators, vitamins, reporter molecules, cholesterol, phospholipids, biotin, phenazine, phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, dyes, antibodies, haptens, antigens, enzymes, and detection reagents including, but not limited to, fluorophores, metals including, but not limited to, gold, metal chelates, chromophores, fluorophore precursors and chromophore precursors, that possess or are modified to possess a hydrazino, oxyamino or carbonyl group that is complementary to the carbonyl, oxyamino or hydrazino group of the oligonucleotide of

formula (II) for formation of hydrazone or oxime linkage. Dendrimeric compounds that possess a plurality of detectable groups, including, but not limited to, reporter molecules, fluorophores, chromophores, fluorophore precursors and chromophore precursors, are also
5 contemplated herein as second components. Fluorophore precursors and chromophore precursors are compounds that react with the hydrazino, oxyamino or carbonyl group of the modified oligonucleotide to form a fluorogenic or chromogenic group for analysis. Such groups are preferred due to the absence of background noise in the resulting assay. Preferred
10 conjugates include those containing a hydrazone linkage.

In embodiments where X is a hydrazino or oxyamino group, the second component can be modified to possess an epoxide, α -bromo-carbonyl, maleimide, maleic anhydride, isothiocyanate or isocyanate group. Such second components can be prepared by methods provided
15 herein or by other methods well known to those of skill in the art. For example, reaction of pentafluorophenyl 4-isothiocyanatobenzoate with an amino or hydroxy group of a second component, such as a protein or oligosaccharide, results in formation of an isothiocyanato modified second component. Reaction of the hydrazino or oxyamino group of the modified
20 oligonucleotide with the epoxide, α -bromocarbonyl, maleimide, maleic anhydride, isothiocyanate or isocyanate group of the second component results in covalent attachment of the oligonucleotide to the second component to form the conjugates provided herein.

In certain embodiments herein, the oligonucleotide conjugates are
25 selected with the proviso that the covalent linkage is not an acyl hydrazone ($-C(O)NHNH=$). In preferred embodiments, the oligonucleotide conjugates are prepared from modified oligonucleotides of formula (II) where P^2 is not a phosphoramidate. In other embodiments, the oligonucleotide conjugates are selected with the proviso that the second
30 component is not a protein, particularly a glycoprotein, more particularly an immunoglobulin.

The modified oligonucleotides provided herein can also be conjugated to ligands such as growth factors, membrane-active bacterial proteins and asialoglycoproteins as delivery strategies for antisense therapy and gene therapy (see, e.g., Perales *et al.* (1994) *Eur. J.*

- 5 *Biochem.* 226:255; Cristiano *et al.* (1996) *Cancer Gene Ther.* 3:4; Hoganson *et al.* (1998) *Hum. Gene Ther.* 9:2565; Gottschalk *et al.* (1995) *Gene Ther.* 2:498; Lu *et al.* (1994) *J. Nucl. Med.* 35:269).

Oligonucleotide-antigen conjugates can be used as immunomodulators in regulating airway eosinophilia in bronchial asthma (Shirota *et al.* (2000) *J.*

- 10 *Immunol.* 164:5575). Oligonucleotide probes conjugated to fluorescent molecules are used in fluorescence *in situ* hybridization (FISH) for chromosome classification and the detection of chromosome aberrations (Pinkel *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:2934).

Oligonucleotide conjugates are also used to study the mechanics of DNA

- 15 hybridization, and to investigate protein-DNA contacts of DNA binding proteins (Lannutti *et al.* (1996) *Biochemistry* 35:9821; Brown *et al.* (1997) *Curr. Opin. Biotechnol.* 8:45).

Importantly, the immobilized oligonucleotides or oligonucleotide conjugates provided herein can be formed under aqueous conditions

- 20 without the need for additional reagents, such as a reducing agent.

Methods for the attachment of hydrazino, oxyamino, or carbonyl modified oligonucleotides of formula (II) to appropriately modified surfaces, as described herein, are provided. The attachment is via a covalent hydrazone or oxime bond formed from the hydrazino, oxyamino, or carbonyl group of the modified oligonucleotide and a complementary carbonyl, oxyamino or hydrazino modified surface provided herein (see, FIGURE 2).

- Methods for the conjugation of oligonucleotide first components to second components, including, but not limited to, biopolymers as defined herein, polymers including, but not limited to, polyamines, polyamides, polyethers and polyethylene glycols, and other compounds of interest
- 30

- herein for use in assays, kits, diagnostic arrays, and the like, including, but not limited to, intercalators, vitamins, reporter molecules, cholesterol, phospholipids, biotin, phenazine, phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, dyes,
- 5 antibodies, haptens, antigens, enzymes, and detection reagents including, but not limited to, fluorophores, metals including, but not limited to, gold, metal chelates, chromophores, fluorophore precursors and chromophore precursors, that possess or are modified to possess a hydrazino, oxyamino or carbonyl group that is complementary to the carbonyl,
- 10 oxyamino or hydrazino group of the oligonucleotide of formula (II) for formation of hydrazone or oxime linkage, are provided. Dendrimeric compounds that possess a plurality of detectable groups, including, but not limited to, reporter molecules, fluorophores, chromophores, fluorophore precursors and chromophore precursors, are also
- 15 contemplated herein as second components. Fluorophore precursors and chromophore precursors are compounds that react with the hydrazino, oxyamino or carbonyl group of the modified oligonucleotide to form a fluorogenic or chromogenic group for analysis.

In particular, methods for conjugation of a hydrazino, oxyamino, or carbonyl modified oligonucleotide of formula (II) with an appropriately

20 modified second component are provided. The conjugation is achieved through a covalent hydrazone or oxime bond formed from the hydrazino, oxyamino or carbonyl group of the modified oligonucleotide and a second component possessing a complementary carbonyl, oxyamino or hydrazino

25 group.

In all embodiments herein, it is preferred that one, more preferably both, of the reactive partners (e.g., the hydrazino, oxyamino or carbonyl groups) are aromatic or heteroaromatic. Thus, in preferred embodiments, the compounds of formula (I) will be aryl or heteroaryl hydrazino or

30 oxyamino derivatives, or aryl or heteroaryl carbonyl derivatives. In more preferred embodiments, the coupling partner (e.g., the modified solid

surface or the second component) will also possess an aryl or heteroaryl hydrazino or oxyamino group, or an aryl or heteroaryl carbonyl group. Hydrazone and oxime linkages formed from these preferred groups are more stable than the corresponding aliphatic hydrazone and oxime linkages, and thus are more preferred in certain applications.

A composition, comprising one or more nucleoside triphosphates and a monomer provided herein. Methods using the composition to synthesize nucleic acid molecules that include the monomers provided herein incorporated into the chains. Nucleic acid molecules (i.e. oligonucleotides) in which the monomers provided herein are at the end are also provided; these are produced by providing the monomers as chain terminators, such as dideoxynucleotides.

DESCRIPTION OF THE FIGURES

FIGURE 1 illustrates incorporation of the capture nucleoside triphosphate (CNTP) monomers provided herein into an RNA/DNA target to form a modified RNA/DNA and capture of the modified RNA/DNA on a solid surface.

FIGURE 2 illustrates hydrazino and oxyamino derivatives provided herein.

FIGURE 3 illustrates carbonyl groups exemplified herein.

FIGURE 4 shows the immobilization of a hydrazino or oxyamino modified oligonucleotide on a carbonyl modified solid support.

FIGURE 5 shows the synthesis of a carbonyl (aldehyde) triphosphate monomer provided herein.

FIGURE 6 shows the synthesis of a hydrazino modified fluorescein provided herein.

FIGURE 7 illustrates the synthesis of a hydrazino triphosphate monomer provided herein.

FIGURE 8 illustrates the synthesis of an oxyamino triphosphate monomer provided herein.

FIGURE 9 illustrates the modified fluorescent groups F* provided herein.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Definitions

5 Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. All patents, patent applications and publications referred to throughout the disclosure herein are incorporated by reference in their entirety. In the event that there are
10 a plurality of definitions for term herein, those in this section prevail.

As used herein, "hydrazino groups" include, but are not limited to, hydrazines, hydrazides, semicarbazides, carbazides, thiosemicarbazides, thiocarbazides, hydrazine carboxylates and carbonic acid hydrazines (see, e.g., FIGURE 2).

15 As used herein, hydrazone linkages include, but are not limited to, hydrazones, acyl hydrazones, semicarbazones, carbazones, thiosemicarbazones, thiocarbazones, hydrazone carboxylates and carbonic acid hydrazones.

As used herein, an oxyamino group has the formula $-O-NH_2$. An
20 oxime has the formula $-O-N=R$.

As used herein, carbonyl-based derivatives include, but are not limited to, ketones and aldehydes.

As used herein, a "triphosphate group" refers to $(HO)P(O)(O^-)-O-P(O)(O^-)-O-P(O)(O^-)-$.

25 As used herein, a protected hydrazino or a protected oxyamino group refers to a hydrazino or oxyamino group that has been derivatized as a salt of the hydrazino or oxyamino group, including but not limited to, mineral acids salts, such as but not limited to hydrochlorides and sulfates, and salts of organic acids, such as but not limited to acetates, lactates,
30 malates, tartrates, citrates, ascorbates, succinates, butyrates, valerates and fumarates, or any amino or hydrazino protecting group known to

those of skill in the art (see, e.g., Greene et al. (1999) Protective Groups in Organic Synthesis (3rd Ed.) (J. Wiley Sons, Inc.)). Preferred amino and hydrazino protecting groups herein include, but are not limited to, amino or hydrazino protecting group useful in the synthesis of

5 oligonucleotides, more preferably monomethoxytrityl (MMT), dimethoxytrityl (DMT), 9-fluorenylmethoxycarbonyl (Fmoc), acetyl, trifluoroacetyl (TFA), benzoyl, or a hydrazone that is cleaved under mild acidic conditions (e.g., 100 mM acetate, pH 4.5-5.5) including, but not limited to, a hydrazone formed from a lower aliphatic aldehyde or ketone,

10 preferably from acetone, cyclohexanone, propanal or 2-butanone.

As used herein, an oligonucleotide is a nucleic acid, including, but not limited to, a ribonucleic acid (RNA), a deoxyribonucleic acid (DNA), and analogs thereof such as a protein nucleic acid (PNA), of any length, including chromosomes and genomic material, such as PCR products or

15 sequencing reaction products, preferably DNA including double and single stranded forms. Single stranded forms of the oligonucleotides are also provided.

As used herein, a conjugate is a compound containing two components covalently linked together. For example, a first component,

20 e.g., an oligonucleotide, is conjugated through a covalent hydrazone linkage to a second component, as defined herein, to form a conjugate.

As used herein, a biopolymer is any compound found in nature, or derivatives thereof, made up of monomeric units. Biopolymers include, but are not limited to, oligonucleotides, peptides, peptide nucleic acids

25 (PNAs), glycoproteins and oligosaccharides. Thus, the monomeric units include, but are not limited to, nucleotides, nucleosides, amino acids, PNA monomers, monosaccharides, and derivatives thereof.

As used herein, a macromolecule refers to a molecule of colloidal size (i.e., of high molecular weight), including, but not limited to,

30 proteins, polynucleic acids, polysaccharides and carbohydrates.

As used herein, complementary reactive groups are those that, when reacted together, form a covalent linkage, including, but not limited to, a hydrazone or oxime linkage. Thus, a hydrazino group, as defined herein, is complementary to a carbonyl group. An oxyamino group is also
5 complementary to a carbonyl group.

As used herein, a reporter molecule refers to a molecule, such as an enzyme or indicator, which is capable of generating a detectable signal (e.g., by colorimetric, chemiluminescent, bioluminescent, fluorescent, or potentiometric means) when contacted with a suitable substrate under
10 appropriate reaction conditions. Exemplary reporter enzymes include, but are not limited to, alkaline phosphatase, horseradish peroxidase, β -galactosidase, aryl esterase, sulfatase and urease.

As used herein, a nucleobase is a heterocyclic moiety that is found in naturally occurring oligonucleotides, including ribonucleic acids (RNA)
15 and deoxyribonucleic acids (DNA), and analogs thereof, including deaza analogs. Preferred nucleobases include, but are not limited to, cytosines, uracils, adenines, guanines and thymines, and analogs thereof including deaza analogs.

As used herein, a derivative of a compound includes a salt, ester,
20 enol ether, enol-ester, solvate or hydrate thereof that can be readily prepared by those of skill in this art using known methods for such derivatization. Salts include, but are not limited to, amine salts, such as but not limited to N,N'-dibenzylethylenediamine, chloroprocaine, choline, ammonia, diethanolamine and other hydroxyalkylamines, ethylenediamine,
25 N-methylglucamine, procaine, N-benzylphenethylamine, 1-para-chloro-benzyl-2-pyrrolidin-1'-ylmethylbenzimidazole, diethylamine and other alkylamines, piperazine and tris(hydroxymethyl)aminomethane; alkali metal salts, such as but not limited to lithium, potassium and sodium; alkali earth metal salts, such as but not limited to barium, calcium and
30 magnesium; transition metal salts, such as but not limited to zinc; and other metal salts, such as but not limited to sodium hydrogen phosphate

and disodium phosphate; and also including, but not limited to, salts of mineral acids, such as but not limited to hydrochlorides and sulfates; and salts of organic acids, such as but not limited to acetates, lactates, malates, tartrates, citrates, ascorbates, succinates, butyrates, valerates and fumarates. Esters include, but are not limited to, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl, heteroaralkyl, cycloalkyl and heterocyclyl esters of acidic groups, including, but not limited to, carboxylic acids, phosphoric acids, phosphinic acids, sulfonic acids, sulfinic acids and boronic acids. Enol ethers include, but are not limited to, derivatives of formula $C=C(OR)$ where R is hydrogen, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl, heteroaralkyl, cycloalkyl or heterocyclyl. Enol esters include, but are not limited to, derivatives of formula $C=C(OC(O)R)$ where R is hydrogen, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl, heteroaralkyl, cycloalkyl or heterocyclyl. Solvates and hydrates are complexes of a compound with one or more solvent or water molecule, preferably 1 to about 100, more preferably 1 to about 10, most preferably one to about 2, 3 or 4, solvent or water molecules.

It is to be understood that the compounds provided herein can contain chiral centers. Such chiral centers can be of either the (R) or (S) configuration, or can be a mixture thereof. Thus, the compounds provided herein can be enantiomerically pure, or be stereoisomeric or diastereomeric mixtures. In the case of amino acid residues, such residues can be of either the L- or D-form. The preferred configuration for naturally occurring sugar residues is D.

As used herein, alkyl, alkenyl and alkynyl carbon chains, if not specified, contain from 1 to 20 carbons, preferably 1 to 16 carbons, and are straight or branched. Alkenyl carbon chains of from 2 to 20 carbons preferably contain 1 to 8 double bonds, and the alkenyl carbon chains of 1 to 16 carbons preferably contain 1 to 5 double bonds. Alkynyl carbon chains of from 2 to 20 carbons preferably contain 1 to 8 triple bonds, and the alkynyl carbon chains of 2 to 16 carbons preferably contain 1 to 5

triple bonds. Exemplary alkyl, alkenyl and alkynyl groups herein include, but are not limited to, methyl, ethyl, propyl, isopropyl, isobutyl, n-butyl, sec-butyl, tert-butyl, isopentyl, neopentyl, tert-pentyl and isohexyl. The alkyl, alkenyl and alkynyl groups, unless otherwise specified, can be optionally substituted, with one or more groups, preferably alkyl group substituents that can be the same or different. As used herein, lower alkyl, lower alkenyl, and lower alkynyl refer to carbon chains having less than about 6 carbons. As used herein, "alk(en)(yn)yl" refers to an alkyl group containing at least one double bond and at least one triple bond.

10 As used herein, an "alkyl group substituent" includes halo, haloalkyl, preferably halo lower alkyl, aryl, hydroxy, alkoxy, aryloxy, alkyloxy, alkylthio, arylthio, aralkyloxy, aralkylthio, carboxy, alkoxycarbonyl, oxo and cycloalkyl.

As used herein, "aryl" refers to cyclic groups containing from 5 to 15 19 carbon atoms. Aryl groups include, but are not limited to groups, such as fluorenyl, substituted fluorenyl, phenyl, substituted phenyl, naphthyl and substituted naphthyl, in which the substituent is lower alkyl, halogen, or lower alkoxy.

As used herein, an "aryl group substituent" includes alkyl, cyclo- 20 alkyl, cycloalkylalkyl, aryl, heteroaryl optionally substituted with 1 or more, preferably 1 to 3, substituents selected from halo, halo alkyl and alkyl, aralkyl, heteroaralkyl, alkenyl containing 1 to 2 double bonds, alkynyl containing 1 to 2 triple bonds, alk(en)(yn)yl groups, halo, pseudohalo, cyano, hydroxy, haloalkyl and polyhaloalkyl, preferably halo 25 lower alkyl, especially trifluoromethyl, formyl, alkylcarbonyl, arylcarbonyl that is optionally substituted with 1 or more, preferably 1 to 3, substituents selected from halo, halo alkyl and alkyl, heteroarylcarbonyl, carboxy, alkoxycarbonyl, aryloxycarbonyl, aminocarbonyl, alkylaminocarbonyl, dialkylaminocarbonyl, arylaminocarbonyl, diarylaminocarbonyl, 30 aralkylaminocarbonyl, alkoxy, aryloxy, perfluoroalkoxy, alkenyloxy, alkynyloxy, arylalkoxy, aminoalkyl, alkylaminoalkyl, dialkylaminoalkyl,

- arylaminoalkyl, amino, alkylamino, dialkylamino, arylamino, alkylarylamino, alkylcarbonylamino, arylcarbonylamino, azido, nitro, mercapto, alkylthio, arylthio, perfluoroalkylthio, thiocyano, isothiocyano, alkylsulfinyl, alkylsulfonyl, arylsulfinyl, arylsulfonyl, aminosulfonyl,
- 5 alkylaminosulfonyl, dialkylaminosulfonyl and arylaminosulfonyl.

As used herein, "aralkyl" refers to an alkyl group in which one of the hydrogen atoms of the alkyl is replaced by an aryl group.

As used herein, "heteroaralkyl" refers to an alkyl group in which one of the hydrogen atoms of the alkyl is replaced by a heteroaryl group.

- 10 As used herein, "cycloalkyl" refers to a saturated mono- or multi-cyclic ring system, preferably of 3 to 10 carbon atoms, more preferably 3 to 6 carbon atoms; cycloalkenyl and cycloalkynyl refer to mono- or multicyclic ring systems that respectively include at least one double bond and at least one triple bond. Cycloalkenyl and cycloalkynyl groups
- 15 can preferably contain 3 to 10 carbon atoms, with cycloalkenyl groups more preferably containing 4 to 7 carbon atoms and cycloalkynyl groups more preferably containing 8 to 10 carbon atoms. The ring systems of the cycloalkyl, cycloalkenyl and cycloalkynyl groups can be composed of one ring or two or more rings which can be joined together in a fused,
- 20 bridged or spiro-connected fashion, and can be optionally substituted with one or more alkyl group substituents. "Cycloalk(en)(yn)yl" refers to a cycloalkyl group containing at least one double bond and at least one triple bond.

- As used herein, "heteroaryl" refers to a monocyclic or multicyclic
- 25 ring system, preferably of about 5 to about 15 members where one or more, more preferably 1 to 3 of the atoms in the ring system is a heteroatom, that is, an element other than carbon, for example, nitrogen, oxygen and sulfur atoms. The heteroaryl can be optionally substituted with one or more, preferably 1 to 3, aryl group substituents. The
- 30 heteroaryl group can be optionally fused to a benzene ring. Exemplary heteroaryl groups include, for example, furyl, imidazynyl, pyrrolidinyl,

pyrimidinyl, tetrazolyl, thienyl, pyridyl, pyrrolyl, N-methylpyrrolyl, quinolinyl and isoquinolinyl, with pyridyl and quinolinyl being preferred.

As used herein, "heterocyclic" refers to a monocyclic or multicyclic ring system, preferably of 3 to 10 members, more preferably 4 to 7
5 members, even more preferably 5 to 6 members, where one or more, preferably 1 to 3 of the atoms in the ring system is a heteroatom, that is, an element other than carbon, for example, nitrogen, oxygen and sulfur atoms. The heterocycle can be optionally substituted with one or more, preferably 1 to 3 aryl group substituents. Preferred substituents of the
10 heterocyclic group include hydroxy, amino, alkoxy containing 1 to 4 carbon atoms, halo lower alkyl, including trihalomethyl, such as trifluoromethyl, and halogen. As used herein, the term heterocycle can include reference to heteroaryl.

As used herein, the nomenclature alkyl, alkoxy, carbonyl, etc. are
15 used as is generally understood by those of skill in this art. For example, as used herein alkyl refers to saturated carbon chains that contain one or more carbons; the chains can be straight or branched or include cyclic portions or be cyclic. As used herein, alicyclic refers to aryl groups that are cyclic.

20 Where the number of any given substituent is not specified (e.g., "haloalkyl"), there can be one or more substituents present. For example, "haloalkyl" can include one or more of the same or different halogens. As another example, "C₁₋₃alkoxyphenyl" can include one or more of the same or different alkoxy groups containing one, two or three carbons.

25 As used herein, "halogen" or "halide" refers to F, Cl, Br or I.

As used herein, pseudohalides are compounds that behave substantially similar to halides. Such compounds can be used in the same manner and treated in the same manner as halides (X, in which X is a halogen, such as Cl or Br). Pseudohalides include, but are not limited
30 to, cyanide, cyanate, thiocyanate, selenocyanate, trifluoromethoxy, trifluoromethyl and azide.

As used herein, "haloalkyl" refers to a lower alkyl radical in which one or more of the hydrogen atoms are replaced by halogen including, but not limited to, chloromethyl, trifluoromethyl, 1-chloro-2-fluoroethyl and the like.

- 5 As used herein, "haloalkoxy" refers to RO- in which R is a haloalkyl group.

As used herein, "sulfinyl" or "thionyl" refers to -S(O)-. As used herein, "sulfonyl" or "sulfuryl" refers to -S(O)₂-. As used herein, "sulfo" refers to -S(O)₃-.

- 10 As used herein, "carboxy" refers to a divalent radical, -C(O)O-.

As used herein, "aminocarbonyl" refers to -C(O)NH₂.

- As used herein, "alkylaminocarbonyl" refers to -C(O)NHR in which R is hydrogen or alkyl, preferably lower alkyl. As used herein "dialkylaminocarbonyl" as used herein refers to -C(O)NR'R in which R' and R are independently selected from hydrogen or alkyl, preferably lower alkyl; "carboxamide" refers to groups of formula -NR'COR.
- 15

As used herein, "diarylaminocarbonyl" refers to -C(O)NRR' in which R and R' are independently selected from aryl, preferably lower aryl, more preferably phenyl.

- 20 As used herein, "aralkylaminocarbonyl" refers to -C(O)NRR' in which one of R and R' is aryl, preferably lower aryl, more preferably phenyl, and the other of R and R' is alkyl, preferably lower alkyl.

As used herein, "arylaminocarbonyl" refers to -C(O)NHR in which R is aryl, preferably lower aryl, more preferably phenyl.

- 25 As used herein, "alkoxycarbonyl" refers to -C(O)OR in which R is alkyl, preferably lower alkyl.

As used herein, "aryloxycarbonyl" refers to -C(O)OR in which R is aryl, preferably lower aryl, more preferably phenyl.

- 30 As used herein, "alkoxy" and "alkylthio" refer to RO- and RS-, in which R is alkyl, preferably lower alkyl.

As used herein, "aryloxy" and "arylthio" refer to RO- and RS-, in which R is aryl, preferably lower aryl, more preferably phenyl.

As used herein, "alkylene" refers to a straight, branched or cyclic, preferably straight or branched, bivalent aliphatic hydrocarbon group, preferably having from 1 to about 20 carbon atoms, more preferably 1 to 12 carbons, even more preferably lower alkylene. The alkylene group is optionally substituted with one or more "alkyl group substituents." There can be optionally inserted along the alkylene group one or more oxygen, sulphur or substituted or unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary alkylene groups include methylene (-CH₂-), ethylene (-CH₂CH₂-), propylene (-(CH₂)₃-), cyclohexylene (-C₆H₁₀-), methylenedioxy (-O-CH₂-O-) and ethylenedioxy (-O-(CH₂)₂-O-). The term "lower alkylene" refers to alkylene groups having 1 to 6 carbons. Preferred alkylene groups are lower alkylene, with alkylene of 1 to 3 carbon atoms being particularly preferred.

As used herein, "alkenylene" refers to a straight, branched or cyclic, preferably straight or branched, bivalent aliphatic hydrocarbon group, preferably having from 2 to about 20 carbon atoms and at least one double bond, more preferably 1 to 12 carbons, even more preferably lower alkenylene. The alkenylene group is optionally substituted with one or more "alkyl group substituents." There can be optionally inserted along the alkenylene group one or more oxygen, sulphur or substituted or unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary alkenylene groups include -CH=CH-CH=CH- and -CH=CH-CH₂-. The term "lower alkenylene" refers to alkenylene groups having 2 to 6 carbons. Preferred alkenylene groups are lower alkenylene, with alkenylene of 3 to 4 carbon atoms being particularly preferred.

As used herein, "alkynylene" refers to a straight, branched or cyclic, preferably straight or branched, bivalent aliphatic hydrocarbon

group, preferably having from 2 to about 20 carbon atoms and at least one triple bond, more preferably 1 to 12 carbons, even more preferably lower alkynylene. The alkynylene group is optionally substituted with one or more "alkyl group substituents." There can be optionally inserted

5 along the alkynylene group one or more oxygen, sulphur or substituted or unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary alkynylene groups include $-\text{C}\equiv\text{C}-\text{C}\equiv\text{C}-$, $-\text{C}\equiv\text{C}-$ and $-\text{C}\equiv\text{C}-\text{CH}_2-$. The term "lower alkynylene" refers to alkynylene groups having 2 to 6 carbons. Preferred alkynylene

10 groups are lower alkynylene, with alkynylene of 3 to 4 carbon atoms being particularly preferred.

As used herein, "alk(en)(yn)ylene" refers to a straight, branched or cyclic, preferably straight or branched, bivalent aliphatic hydrocarbon group, preferably having from 2 to about 20 carbon atoms and at least

15 one triple bond, and at least one double bond; more preferably 1 to 12 carbons, even more preferably lower alk(en)(yn)ylene. The alk(en)(yn)ylene group is optionally substituted with one or more "alkyl group substituents." There can be optionally inserted along the alkynylene group one or more oxygen, sulphur or substituted or

20 unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary alk(en)(yn)ylene groups include $-\text{C}=\text{C}-(\text{CH}_2)_n-\text{C}\equiv\text{C}-$, where n is 1 or 2. The term "lower alk(en)(yn)ylene" refers to alk(en)(yn)ylene groups having up to 6 carbons. Preferred alk(en)(yn)ylene groups are lower alk(en)(yn)ylene,

25 with alk(en)(yn)ylene of 4 carbon atoms being particularly preferred.

As used herein, "arylene" refers to a monocyclic or polycyclic, preferably monocyclic, bivalent aromatic group, preferably having from 5 to about 20 carbon atoms and at least one aromatic ring, more preferably

5 to 12 carbons, even more preferably lower arylene. The arylene group

30 is optionally substituted with one or more "alkyl group substituents." There can be optionally inserted around the arylene group one or more

oxygen, sulphur or substituted or unsubstituted nitrogen atoms, where the nitrogen substituent is alkyl as previously described. Exemplary arylene groups include 1,2-, 1,3- and 1,4-phenylene. The term "lower arylene" refers to arylene groups having 5 or 6 carbons. Preferred
 5 arylene groups are lower arylene.

As used herein, "heteroarylene" refers to a bivalent monocyclic or multicyclic ring system, preferably of about 5 to about 15 members where one or more, more preferably 1 to 3 of the atoms in the ring system is a heteroatom, that is, an element other than carbon, for
 10 example, nitrogen, oxygen and sulfur atoms. The heteroarylene group can be optionally substituted with one or more, preferably 1 to 3, aryl group substituents.

As used herein, "alkylidene" refers to a bivalent group, such as $=CR'R''$, which is attached to one atom of another group, forming a
 15 double bond. Exemplary alkylidene groups are methylenidene ($=CH_2$) and ethylenidene ($=CHCH_3$). As used herein, "aralkylidene" refers to an alkylidene group in which either R' or R'' is an aryl group.

As used herein, "amido" refers to the bivalent group $-C(O)NH-$. "Thioamido" refers to the bivalent group $-C(S)NH-$. "Oxyamido" refers to
 20 the bivalent group $-OC(O)NH-$. "Thioamido" refers to the bivalent group $-SC(O)NH-$. "Dithioamido" refers to the bivalent group $-SC(S)NH-$. "Ureido" refers to the bivalent group $-HNC(O)NH-$. "Thioureido" refers to the bivalent group $-HNC(S)NH-$.

As used herein, "semicarbazide" refers to $-NHC(O)NHNH-$.
 25 "Carbazate" refers to the bivalent group $-OC(O)NHNH-$. "Isothiocarbazate" refers to the bivalent group $-SC(O)NHNH-$. "Thiocarbazate" refers to the bivalent group $-OC(S)NHNH-$. "Sulfonylhydrazide" refers to the group $-SO_2NHNH-$. "Hydrazide" refers to the bivalent group $-C(O)NHNH-$. "Azo" refers to the bivalent group $-N=N-$.
 30 "Hydrazinyl" refers to the bivalent group $-NH-NH-$.

As used herein, the term "amino acid" refers to α -amino acids which are racemic, or of either the D- or L-configuration. The designation "d" preceding an amino acid designation (e.g., dAla, dSer, dVal, etc.) refers to the D-isomer of the amino acid. The designation "dl" preceding
5 an amino acid designation (e.g., dlPip) refers to a mixture of the L- and D-isomers of the amino acid.

As used herein, when any particular group, such as phenyl or pyridyl, is specified, this means that the group is unsubstituted or is substituted. Preferred substituents where not specified are halo, halo,
10 lower alkyl, and lower alkyl.

As used herein, a composition refers to any mixture of two or more products or compounds. It can be a solution, a suspension, liquid, powder, a paste, aqueous, non-aqueous or any combination thereof.

As used herein, a combination refers to any association between
15 two or more items.

As used herein, fluid refers to any composition that can flow. Fluids thus encompass compositions that are in the form of semi-solids, pastes, solutions, aqueous mixtures, gels, lotions, creams and other such compositions.

20 As used herein, substantially identical to a product means sufficiently similar so that the property of interest is sufficiently unchanged so that the substantially identical product can be used in place of the product.

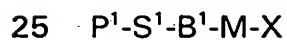
As used herein, the abbreviations for any protective groups, amino
25 acids and other compounds, are, unless indicated otherwise, in accord with their common usage, recognized abbreviations, or the IUPAC-IUB Commission on Biochemical Nomenclature (see, Biochem. 1972, 11, 942).

A. Nucleoside Triphosphate Monomers

Nucleoside triphosphate monomers are provided. The monomers are incorporated into oligonucleotide first components during enzymatic synthesis. The resulting modified oligonucleotides are readily immobilized
 5 or conjugated through a hydrazone or oxime linkage. The reagents possess a triphosphate-ribose-nucleobase motif for recognition by the enzymatic catalyst.

The monomers also possess, in addition to the triphosphate-ribose-nucleobase motif, a protected or unprotected hydrazino, protected or
 10 unprotected oxyamino, or a carbonyl moiety for formation of a hydrazone or oxime linkage with an appropriately modified surface or second component. The hydrazino moiety can be an aliphatic, aromatic or heteroaromatic hydrazine, semicarbazide, carbazide, hydrazide, thiosemicarbazide, thiocarazide, carbonic acid dihydrazine or hydrazine
 15 carboxylate (see, FIGURE 2). The protecting groups are salts of the hydrazino or oxyamino group, or any amino or hydrazino protecting groups known to those of skill in the art (see, e.g., Greene et al. (1999) Protective Groups in Organic Synthesis (3rd Ed.) (J. Wiley Sons, Inc.)). The carbonyl moiety can be any carbonyl containing group capable of
 20 forming a hydrazone or oxime linkage with one or more of the above hydrazino or oxyamino moieties. Preferred carbonyl moieties include aldehydes and ketones.

In one embodiment, the monomers for use in the methods provided herein have formula (I):



or a derivative thereof, as defined herein, where P^1 is a triphosphate group, as defined herein, and is preferably at the 5' position of S^1 ; X is a protected or unprotected hydrazino group; S^1 is a ribose, deoxyribose including a 2'-deoxyribose, or a dideoxyribose including a 2',3'-
 30 dideoxyribose; B^1 is a nucleobase; and M is a divalent group having any combination, preferably 1-2000, more preferably 1-1000, particularly 1-

- 100, most preferably 1-50 or 1-25 or 1-10, of the following groups, which can be combined in any order: arylene, heteroarylene, cycloalkylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^1)$, $P(D)_b(ER^1)$, $N(R^1)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^1 ; D is S or O; and E is S, O or NR^1 ; each R^1 is a monovalent group independently selected from hydrogen and M^1-R^4 ; each M^1 is a divalent group independently having any combination of the following groups, which groups can be combined in any order: a direct link, arylene, heteroarylene, cycloalkylene, $C(R^5)_2$, $-C(R^5)=C(R^5)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^5)$, $P(D)_b(ER^5)$, $N(R^5)$, $N(COR^5)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^5 ; D is S or O; and E is S, O or NR^5 ; R^4 and R^5 are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, $SiR^6R^7R^8$, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; R^9 and R^{10} are each independently selected from hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heteroaryl, heteroaralkyl and heterocyclyl; R^2 and R^3 are selected from (i) or (ii) as follows: (i) R^2 and R^3 are independently selected from among hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl and heteroaryl; or (ii) R^2 and R^3 together form alkylene, alkenylene or cycloalkylene; R^6 , R^7 and R^8 are each independently a monovalent group selected from hydrogen, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; and R^1 , R^2 , R^3 , R^4 , R^5 , R^6 , R^7 , R^8 , R^9 and R^{10} can be substituted with one or more substituents each independently selected from Z, wherein Z is selected from alkyl, alkenyl, alkynyl, aryl, cycloalkyl, cycloalkenyl,

hydroxy, $S(O)_hR^{20}$, $NR^{20}R^{21}$, $COOR^{20}$, COR^{20} , $CONR^{20}R^{21}$, $OC(O)NR^{20}R^{21}$, $N(R^{20})C(O)R^{21}$, alkoxy, aryloxy, heteroaryl, heterocyclyl, heteroaryloxy, heterocyclyloxy, aralkyl, aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, aralkoxy, heteroaralkoxy, alkoxycarbonyl, carbamoyl, thiocarbamoyl, alkoxycarbonyl, carboxyaryl, halo, pseudohalo, haloalkyl and carboxamido; h is 0, 1 or 2; and R^{20} and R^{21} are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, trialkylsilyl, dialkylarylsilyl, alkyl diarylsilyl, triarylsilyl, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy, amino, amido, alkylamino, dialkylamino, alkylaryl amino, diarylamino and arylamino.

In another embodiment, the monomers for use in the methods provided herein have formula (I):

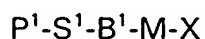
$$P^1-S^1-B^1-M-X$$

or a derivative thereof, as defined herein, where P^1 is a triphosphate group, as defined herein, and is preferably at the 5' position of S^1 ; X is a protected or unprotected oxyamino group; S^1 is a ribose, a deoxyribose including a 2'-deoxyribose, or a dideoxyribose including a 2',3'-dideoxyribose; B^1 is a nucleobase; and M is a divalent group having any combination, preferably 1-2000, more preferably 1-1000, particularly 1-100, most preferably 1-50 or 1-25 or 1-10, of the following groups, which can be combined in any order: arylene, heteroarylene, cycloalkylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^1)$, $P(D)_b(ER^1)$, $N(R^1)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^1 ; D is S or O; and E is S, O or NR^1 ; each R^1 is a monovalent group independently selected from hydrogen and M^1-R^4 ; each M^1 is a divalent group independently having any combination of the following groups, which groups can be combined in any order: a direct link, arylene, heteroarylene, cycloalkylene, $C(R^5)_2$, $-C(R^5)=C(R^5)-$,

- $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^5)$, $P(D)_b(ER^5)$, $N(R^5)$, $N(COR^5)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^5 ; D is S or O; and E is S, O or NR^5 ; R^4 and R^5 are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, $SiR^6R^7R^8$, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; R^9 and R^{10} are each independently selected from hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heteroaryl, heteroaralkyl and heterocyclyl; R^2 and R^3 are selected from (i) or (ii) as follows: (i) R^2 and R^3 are independently selected from among hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl and heteroaryl; or (ii) R^2 and R^3 together form alkylene, alkenylene or cycloalkylene; R^6 , R^7 and R^8 are each independently a monovalent group selected from hydrogen, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; and
- R^1 , R^2 , R^3 , R^4 , R^5 , R^6 , R^7 , R^8 , R^9 and R^{10} can be substituted with one or more substituents each independently selected from Z, wherein Z is selected from alkyl, alkenyl, alkynyl, aryl, cycloalkyl, cycloalkenyl, hydroxy, $S(O)_hR^{20}$, $NR^{20}R^{21}$, $COOR^{20}$, COR^{20} , $CONR^{20}R^{21}$, $OC(O)NR^{20}R^{21}$, $N(R^{20})C(O)R^{21}$, alkoxy, aryloxy, heteroaryl, heterocyclyl, heteroaryloxy, heterocycliloxy, aralkyl, aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, aralkoxy, heteroaralkoxy, alkoxycarbonyl, carbamoyl, thiocarbamoyl, alkoxycarbonyl, carboxyaryl, halo, pseudohalo, haloalkyl and carboxamido; h is 0, 1 or 2; and R^{20} and R^{21} are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, trialkylsilyl, dialkylarylsilyl, alkyl diarylsilyl, triarylsilyl, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl,

heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy, amino, amido, alkylamino, dialkylamino, alkylaryl amino, diarylamino and arylamino.

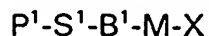
- 5 In another embodiment, the monomers for use in the methods provided herein have formula (I):



- or a derivative thereof, as defined herein, where P^1 is a triphosphate group, as defined herein, and is preferably at the 5' position of S^1 ; X is a
- 10 carbonyl group; S^1 is a ribose, a deoxyribose including 2'-deoxyribose, or a dideoxyribose including 2',3'-dideoxyribose; B^1 is a nucleobase; and M is a divalent group having any combination, preferably 1-2000, more preferably 1-1000, particularly 1-100, most preferably 1-50 or 1-25 or 1-10, of the following groups, which can be combined in any order:
- 15 arylene, heteroarylene, cycloalkylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^1)$, $P(D)_b(ER^1)$, $N(R^1)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^1 ; D is S or O; and E is S, O or NR^1 ; each R^1 is a monovalent group independently selected from hydrogen and M^1-R^4 ; each M^1 is a divalent
- 20 group independently having any combination of the following groups, which groups can be combined in any order: a direct link, arylene, heteroarylene, cycloalkylene, $C(R^5)_2$, $-C(R^5)=C(R^5)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^5)$, $P(D)_b(ER^5)$, $N(R^5)$, $N(COR^5)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or
- 25 NR^5 ; D is S or O; and E is S, O or NR^5 ; R^4 and R^5 are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, $SiR^6R^7R^8$, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl,
- 30 heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; R^9 and R^{10} are each independently selected from hydrogen,

- alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heteroaryl, heteroaralkyl and heterocyclyl; R^2 and R^3 are selected from (i) or (ii) as follows: (i) R^2 and R^3 are independently selected from among hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl and heteroaryl; or (ii) R^2 and R^3 together form
- 5 alkylene, alkenylene or cycloalkylene; R^6 , R^7 and R^8 are each independently a monovalent group selected from hydrogen, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy,
- 10 alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; and
- R^1 , R^2 , R^3 , R^4 , R^5 , R^6 , R^7 , R^8 , R^9 and R^{10} can be substituted with one or more substituents each independently selected from Z, wherein Z is selected from alkyl, alkenyl, alkynyl, aryl, cycloalkyl, cycloalkenyl, hydroxy, $S(O)_hR^{20}$, $NR^{20}R^{21}$, $COOR^{20}$, COR^{20} , $CONR^{20}R^{21}$, $OC(O)NR^{20}R^{21}$,
- 15 $N(R^{20})C(O)R^{21}$, alkoxy, aryloxy, heteroaryl, heterocyclyl, heteroaryloxy, heterocyclyloxy, aralkyl, aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, aralkoxy, heteroaralkoxy, alkoxycarbonyl, carbamoyl, thiocarbamoyl, alkoxycarbonyl, carboxyaryl, halo, pseudohalo, haloalkyl and carboxamido; h is 0, 1 or 2; and R^{20} and R^{21} are each
- 20 independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, trialkylsilyl, dialkylarylsilyl, alkyl diarylsilyl, triarylsilyl, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy,
- 25 alkoxy, aryloxy, aralkoxy, heteroaralkoxy, amino, amido, alkylamino, dialkylamino, alkylarylamino, diarylamino and arylamino.

In certain embodiments, the reagents for use in the methods provided herein are triphosphates that have formula (I):



- 30 or a derivative thereof, wherein S^1 is ribose, 2'-deoxyribose, or 2',3'-dideoxyribose; X is a protected or unprotected hydrazino, protected or

unprotected oxyamino, or a carbonyl group; B¹ is a cytosine, uracil, adenine, guanine or thymine, or an analog thereof, including, but not limited to, 7-deazaadenine, 7-deazaguanine and 7-deazahypoxanthine; and P¹ and M are selected as above.

- 5 M preferably has 1-2000, more preferably 1-1000, particularly 1-100, most preferably 1-50 or 1-25 or 1-10 of the groups as defined above. In certain embodiments, M has 1-50, preferably 1-25, more preferably 1-10 of the following groups, which can be combined in any order: arylene, heteroarylene, C(R¹)₂, -C(R¹)=C(R¹)-, -C≡C-, O, S(A)_a,
- 10 N(R¹), N(COR¹) and C(E); where a is 0, 1 or 2; A is O or NR¹; and E is S, O or NR¹; where R¹ is as defined above. In more preferred embodiments, M has 1-10 of the following groups, which can be combined in any order: heteroarylene, C(R¹)₂, -C≡C-, N(R¹) and C(E); where E is S, O or NR¹; where R¹ is as defined above. In other embodiments, M has 1-10 of the
- 15 following groups, which can be combined in any order: arylene, C(R¹)₂, -C≡C-, N(R¹) and C(E); where E is S, O or NR¹; where R¹ is as defined above. R⁵ is preferably hydrogen; and E is preferably oxygen. In further embodiments, M has 1-10 of the following groups, which can be combined in any order: C(R¹)₂, -C≡C-, N(R¹) and C(E); where E is S, O or
- 20 NR¹; where R¹ is as defined above. In these embodiments, the heteroarylene groups of M are preferably monocyclic or fused bicyclic divalent groups, more preferably monocyclic divalent groups, including, but not limited to, thienylene, furylene, pyrrolylene including 2,3-, 2,4- and 2,5-pyrrolylene, pyridylene including 2,3-, 2,4-, 2,5- and 2,6-
- 25 pyridylene, triazinylene including 1,3,5-triazinylene, triazolylene including 1,2,4- and 1,2,3-triazolylene or pyrimidylene including 2,4-, 2,5- and 2,6-pyrimidylene groups. The arylene groups of M are preferably divalent monocyclic, bicyclic or tricyclic groups, including, but not limited to, 1,2-, 1,3- and 1,4-phenylene, and naphthylene.
- 30 In other embodiments, M is a chain of repeating monomer units, preferably 1-2000 monomer units, more preferably 1-1000, particularly 1-

100, most preferably 1-50 or 1-25 or 1-10, including, but not limited to, ethylene oxide, propylene oxide, methacrylamide, or ethylene glycol. Thus, in certain embodiments, M is a polyethylene glycol (PEG), polypropylene glycol (PPG) or acrylate chain. In particular embodiments, M is
 5 a diethyleneglycol, triethyleneglycol or tetraethyleneglycol moiety.

In other embodiments, M contains a cleavable linkage, including, but not limited to, a disulfide, an ester, an enzyme specific peptide, a photocleavable linkage, such as nitroaromatic group, or an acid labile group. Monomers of formula (I) and modified oligonucleotides of formula
 10 (II) containing such M groups have applications in capture assay methods where the analyte is captured by the modified oligonucleotide, the conjugate isolated, and the M group cleaved to provide an isolated analyte. See, e.g., Niemeyer *et al.* (1994) *Nucl. Acids Res.* 22(25):5530-5539.

15 X is preferably $-C(O)R^{30}$, $-Y-N(R^{31})-Y^1-N(R^{32})-Y^2$ or $-O-N(R^{30})-Y^2$; where R^{30} , R^{31} and R^{32} are each independently hydrogen, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl, aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl or cycloalkyl; Y and Y^1 are selected as in (i) or (ii) as follows:

20 (i) Y is a direct link, and Y^1 is a direct link, $C(O)N(R^{35})$, $N(R^{35})C(O)N(R^{36})$, $C(S)N(R^{35})$, $N(R^{35})C(S)N(R^{36})$ or $C(O)N(R^{35})N(R^{36})C(O)N(R^{37})$; or

(ii) Y is $C(O)$ or $OC(O)$, and Y^1 is a direct link; where R^{35} , R^{36} and R^{37} are each independently selected from
 25 among hydrogen, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl, aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl and cycloalkyl; and

Y^2 is a salt of the hydrazino or oxyamino group, including but not limited to, mineral acids salts, such as but not limited to hydrochlorides
 30 and sulfates, and salts of organic acids, such as but not limited to acetates, lactates, malates, tartrates, citrates, ascorbates, succinates,

butyrates, valerates and fumarates, or any amino or hydrazino protecting group (see, e.g., Greene et al. (1999) Protective Groups in Organic Synthesis (3rd Ed.) (J. Wiley Sons, Inc.));

- where R^{30} , R^{31} , R^{32} , R^{35} , R^{36} , R^{37} and Y^2 are unsubstituted or
- 5 substituted with one or more substituents each independently selected from Z, wherein Z is selected from alkyl, alkenyl, alkynyl, aryl, cycloalkyl, cycloalkenyl, hydroxy, $S(O)_h R^{20}$, $NR^{20}R^{21}$, $COOR^{20}$, COR^{20} , $CONR^{20}R^{21}$, $OC(O)NR^{20}R^{21}$, $N(R^{20})C(O)R^{21}$, alkoxy, aryloxy, heteroaryl, heterocyclyl, heteroaryloxy, heterocyclyloxy, aralkyl, aralkenyl, aralkynyl, heteroaralkyl,
 - 10 heteroaralkenyl, heteroaralkynyl, aralkoxy, heteroaralkoxy, alkoxycarbonyl, carbamoyl, thiocarbamoyl, alkoxycarbonyl, carboxyaryl, halo, pseudohalo, haloalkyl and carboxamido; h is 0, 1 or 2; and R^{20} and R^{21} are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, trialkylsilyl, dialkylarylsilyl, alkylidiarylsilyl,
 - 15 triarylsilyl, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy, amino, amido, alkylamino, dialkylamino, alkylarylamino, diarylamino and
 - 20 arylamino.

- In certain embodiments herein, Y^2 is selected from any amino or hydrazino protecting group, preferably monomethoxytrityl (MMT), 9-fluorenylmethoxycarbonyl (Fmoc), acetyl, benzoyl, or a hydrazone or oxime that is cleaved under mild acidic conditions (e.g., 100 mM acetate,
- 25 pH 4.5-5.5) including, but not limited to, a hydrazone or oxime formed from a lower aliphatic aldehyde or ketone, preferably from acetone, cyclohexanone, propanal or 2-butanone. In other embodiments, particularly when X is a protected oxyamino group, the protecting group is a phthalimidyl group.

- 30 The monomers provided herein are added to the growing oligonucleotide chain during enzymatic synthesis. In the case of terminal

modification, wherein purification of the final product using a lipophilic handle is required, the amino or hydrazino protecting group is MMT, or a similar moiety. The final protected hydrazino or protected oxyamino substituted product is deprotected following high performance liquid chromatography (HPLC) purification.

1. Hydrazino Derivatives

In one preferred embodiment, the monomers for use in the methods provided herein contain a triphosphate group, as defined herein, a sugar group, a nucleobase, and a hydrazino group (see, FIGURE 2).

Thus, in this embodiment, the compounds of formula (I) have formula (III):



or a derivative thereof, where P^1 is a triphosphate group, preferably located at the 5' position of S^1 ;

M is a divalent group having any combination, preferably 1-2000, more preferably 1-1000, particularly 1-100, most preferably 1-50 or 1-25 or 1-10, of the following groups, which can be combined in any order:

arylene, heteroarylene, cycloalkylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$,

$>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O , $S(A)_a$, $P(D)_b(R^1)$, $P(D)_b(ER^1)$, $N(R^1)$,

$>N^+(R^2)(R^3)$ and $C(E)$; where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^1 ; D is S or O ; and E is S , O or NR^1 ; each R^1 is a monovalent group

independently selected from hydrogen and M^1-R^4 ; each M^1 is a divalent group independently having any combination of the following groups,

which groups can be combined in any order: a direct link, arylene,

heteroarylene, cycloalkylene, $C(R^5)_2$, $-C(R^5)=C(R^5)-$, $>C=C(R^2)(R^3)$,

$>C(R^2)(R^3)$, $-C\equiv C-$, O , $S(A)_a$, $P(D)_b(R^5)$, $P(D)_b(ER^5)$, $N(R^5)$, $N(COR^5)$,

$>N^+(R^2)(R^3)$ and $C(E)$; where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^5 ; D is S or O ; and E is S , O or NR^5 ; R^4 and R^5 are each independently

selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro,

$SiR^6R^7R^8$, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl,

aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl,

heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; R^9 and R^{10} are each independently selected from hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heteroaryl, heteroaralkyl and heterocyclyl; R^2 and R^3 are selected from (i) or (ii) as follows: (i) R^2 and R^3 are independently selected from among hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl and heteroaryl; or (ii) R^2 and R^3 together form alkylene, alkenylene or cycloalkylene; R^6 , R^7 and R^8 are each independently a monovalent group selected from hydrogen, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ;

Y and Y^1 are selected as in (i) or (ii) as follows:

(i) Y is a direct link, and Y^1 is a direct link, $\text{C}(\text{O})\text{N}(\text{R}^{35})$, $\text{N}(\text{R}^{35})\text{C}(\text{O})\text{N}(\text{R}^{36})$, $\text{C}(\text{S})\text{N}(\text{R}^{35})$, $\text{N}(\text{R}^{35})\text{C}(\text{S})\text{N}(\text{R}^{36})$ or $\text{C}(\text{O})\text{N}(\text{R}^{35})\text{N}(\text{R}^{36})\text{C}(\text{O})\text{N}(\text{R}^{37})$; or

(ii) Y is $\text{C}(\text{O})$ or $\text{OC}(\text{O})$, and Y^1 is a direct link;

where R^{35} , R^{36} and R^{37} are each independently selected from among hydrogen, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl, aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl and cycloalkyl;

Y^2 is a salt of the hydrazino group, including but not limited to, mineral acids salts, such as but not limited to hydrochlorides and sulfates, and salts of organic acids, such as but not limited to acetates, lactates, malates, tartrates, citrates, ascorbates, succinates, butyrates, valerates and fumarates, or any amino or hydrazino protecting group, preferably an amino or hydrazino protecting group useful in the synthesis of oligonucleotides, more preferably monomethoxytrityl (MMT), dimethoxytrityl (DMT), 9-fluorenylmethoxycarbonyl (Fmoc), acetyl, trifluoroacetyl (TFA), benzoyl, or a hydrazone that is cleaved under mild

acidic conditions (e.g., 100 mM acetate, pH 4.5-5.5) including, but not limited to, a hydrazone formed from a lower aliphatic aldehyde or ketone, preferably from acetone, cyclohexanone, propanal or 2-butanone; and

- R^{31} and R^{32} are each independently hydrogen, alkyl, alkenyl,
 5 alkynyl, aryl, heteroaryl, aralkyl, aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl or cycloalkyl;

where R^1 , R^2 , R^3 , R^4 , R^5 , R^6 , R^7 , R^8 , R^9 , R^{10} , R^{31} , R^{32} , R^{35} , R^{36} , R^{37} and Y^2 are unsubstituted or substituted with one or more substituents each independently selected from Z, as defined above.

- 10 In preferred embodiments, R^{31} , R^{32} , R^{35} , R^{36} and R^{37} are hydrogen or lower alkyl, more preferably hydrogen. In these embodiments, the compounds of formula (III) have formula (IV):



- or a derivative thereof, where Y and Y^1 are selected from (i) or (ii) as
 15 follows:

(i) Y is a direct link, and Y^1 is a direct link, C(O)NH, NHC(O)NH, C(S)NH, NHC(S)NH or C(O)NHNHC(O)NH; or

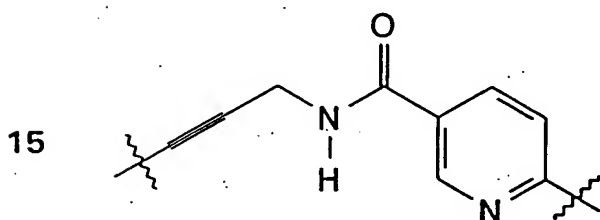
(ii) Y is C(O) or OC(O), and Y^1 is a direct link;
 and P^1 , S^1 , B^1 , M and Y^2 are selected as above.

- 20 In a preferred embodiment, the compounds for use in the methods provided herein are of formula (IV) where M has 1-50, preferably 1-25, more preferably 1-10 of the following groups, which can be combined in any order: arylene, heteroarylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$, $-C\equiv C-$, O, $S(A)_a$, $N(R^1)$, $N(COR^1)$ and C(E); where a is 0, 1 or 2; A is O or NR^1 ; and E
 25 is S, O or NR^1 ; where R^1 is selected as above. In more preferred embodiments, M has 1-10 of the following groups, which can be combined in any order: heteroarylene, $C(R^1)_2$, $-C\equiv C-$, $N(R^1)$ and C(E); where E is S, O or NR^1 ; where R^1 is selected as above. R^1 is preferably hydrogen; and E is preferably oxygen.

- 30 In certain embodiments herein, the group attached to Y is a heteroarylene group, preferably a nitrogen-containing heteroarylene group,

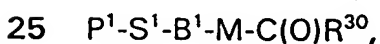
preferably a pyrrolylene including 2,3-, 2,4- and 2,5-pyrrolylene, pyridylene including 2,3-, 2,4-, 2,5- and 2,6-pyridylene, triazinylene including 1,3,5-triazinylene, triazolylene including 1,2,4- and 1,2,3-triazolylene or pyrimidylene including 2,4-, 2,5- and 2,6-pyrimidylene group, more preferably a pyridylene group, most preferably a 2,5-pyridylene group. Thus, in these more preferred embodiments, M includes 2,5-pyridylene attached to Y, preferably Y is at the 2-position of the pyridylene group, and 1-9 of the following groups, which can be combined in any order: CH_2 , $-\text{C}\equiv\text{C}-$, NH and $\text{C}(\text{O})$.

In more preferred embodiments, M has the formula:



2. Carbonyl Derivatives

In another preferred embodiment, the monomers for use in the methods provided herein contain a triphosphate group, as defined herein, a sugar group, a nucleobase, and a carbonyl group. Thus, in this embodiment, the compounds of formula (I) have formula (V):



or a derivative thereof, where P^1 is a triphosphate group, preferably located at the 5' position of S^1 ;

M is a divalent group having any combination, preferably 1-2000, more preferably 1-1000, particularly 1-100, most preferably 1-50 or 1-25

or 1-10, of the following groups, which can be combined in any order: arylene, heteroarylene, cycloalkylene, $\text{C}(\text{R}^1)_2$, $-\text{C}(\text{R}^1)=\text{C}(\text{R}^1)-$, $>\text{C}=\text{C}(\text{R}^2)(\text{R}^3)$, $>\text{C}(\text{R}^2)(\text{R}^3)$, $-\text{C}\equiv\text{C}-$, O, $\text{S}(\text{A})_a$, $\text{P}(\text{D})_b(\text{R}^1)$, $\text{P}(\text{D})_b(\text{ER}^1)$, $\text{N}(\text{R}^1)$, $>\text{N}^+(\text{R}^2)(\text{R}^3)$ and $\text{C}(\text{E})$; where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^1 ; D is S or O; and E is S, O or NR^1 ; each R^1 is a monovalent group

- independently selected from hydrogen and M^1-R^4 ; each M^1 is a divalent group independently having any combination of the following groups, which groups can be combined in any order: a direct link, arylene, heteroarylene, cycloalkylene, $C(R^5)_2$, $-C(R^5)=C(R^5)-$, $>C=C(R^2)(R^3)$,
5 $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^5)$, $P(D)_b(ER^5)$, $N(R^5)$, $N(COR^5)$, $>N^+(R^2)(R^3)$ and $C(E)$; where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^5 ; D is S or O; and E is S, O or NR^5 ; R^4 and R^5 are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, $SiR^6R^7R^8$, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl,
10 aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; R^9 and R^{10} are each independently selected from hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heteroaryl, heteroaralkyl
15 and heterocyclyl; R^2 and R^3 are selected from (i) or (ii) as follows: (i) R^2 and R^3 are independently selected from among hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl and heteroaryl; or (ii) R^2 and R^3 together form alkylene, alkenylene or cycloalkylene; R^6 , R^7 and R^8 are each independently a monovalent group selected from hydrogen, alkyl, alkenyl,
20 alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ;
 R^{30} is hydrogen, alkyl, alkenyl, alkynyl, aryl, heteroaryl, aralkyl,
25 aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl or cycloalkyl;
where R^1 , R^2 , R^3 , R^4 , R^5 , R^6 , R^7 , R^8 , R^9 , R^{10} and R^{30} are unsubstituted or substituted with one or more substituents each independently selected from Z, as defined above.

In preferred embodiments, R^{30} is hydrogen or lower alkyl, more preferably hydrogen or methyl, most preferably hydrogen. In these embodiments, the compounds of formula (V) have formulae (VI):

$P^1-S^1-B^1-M-CHO$ or $P^1-S^1-B^1-M-C(O)Me$,

5 or a derivative thereof, where P^1 , S^1 , B^1 and M are selected as above.

In a preferred embodiment, the compounds for use in the methods provided herein are of formula (VI) where M has 1-50, preferably 1-25, more preferably 1-10 of the following groups, which can be combined in any order: arylene, heteroarylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$, $-C\equiv C-$, O,

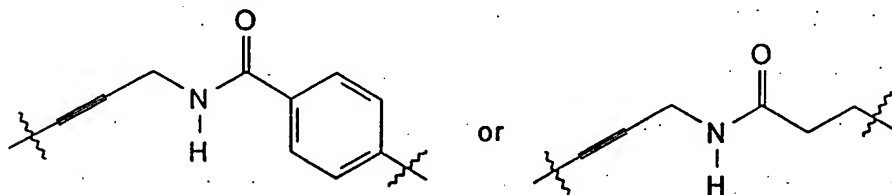
10 $S(A)_a$, $N(R^1)$, $N(COR^1)$ and $C(E)$; where a is 0, 1 or 2; A is O or NR^1 ; and E is S, O or NR^1 ; where R^1 is selected as above. In more preferred embodiments, M has 1-10 of the following groups, which can be combined in any order: arylene, $C(R^1)_2$, $-C\equiv C-$, $N(R^1)$ and $C(E)$; where E is S, O or NR^1 ; where R^1 is selected as above. R^1 is preferably hydrogen;

15 and E is preferably oxygen. In certain embodiments herein, the group attached to CHO or $C(O)Me$ is a arylene group, preferably a monocyclic arylene group, more preferably phenylene. Thus, in these more preferred embodiments, M includes phenylene attached to CHO or $C(O)Me$, and 1-9 of the following groups, which can be combined in any order: CH_2 ,

20 $-C\equiv C-$, NH and $C(O)$. In other embodiments herein, M preferably includes 1-10, more preferably 1-6, of the following groups, which can be combined in any order: CH_2 , $-C\equiv C-$, NH and $C(O)$.

In more preferred embodiments, M has the formulae:

25



30

3. Oxyamino Derivatives

In one preferred embodiment, the reagents for use in the methods provided herein contain a triphosphate group, as defined herein, a sugar group, a nucleobase, and an oxyamino group. The oxyamino group is preferably protected as described above. Thus, in this embodiment, the compounds of formula (I) have formula (VII):



or a derivative thereof, where P^1 is a triphosphate group, preferably located at the 5' position of S^1 ; S^1 and B^1 are selected as above; and

- 10 M is a divalent group having any combination, preferably 1-2000, more preferably 1-1000, particularly 1-100, most preferably 1-50 or 1-25 or 1-10, of the following groups, which can be combined in any order: arylene, heteroarylene, cycloalkylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O , $S(A)_a$, $P(D)_b(R^1)$, $P(D)_b(ER^1)$, $N(R^1)$, $>N^+(R^2)(R^3)$ and $C(E)$; where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^1 ; D is S or O ; and E is S , O or NR^1 ; each R^1 is a monovalent group independently selected from hydrogen and M^1-R^4 ; each M^1 is a divalent group independently having any combination of the following groups, which groups can be combined in any order: a direct link, arylene, heteroarylene, cycloalkylene, $C(R^5)_2$, $-C(R^5)=C(R^5)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O , $S(A)_a$, $P(D)_b(R^5)$, $P(D)_b(ER^5)$, $N(R^5)$, $N(COR^5)$, $>N^+(R^2)(R^3)$ and $C(E)$; where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^5 ; D is S or O ; and E is S , O or NR^5 ; R^4 and R^5 are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, $SiR^6R^7R^8$, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; R^9 and R^{10} are each independently selected from hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heteroaryl, heteroaralkyl, and heterocyclyl; R^2 and R^3 are selected from (i) or (ii) as follows: (i) R^2
- 15
- 20
- 25
- 30

and R^3 are independently selected from among hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl and heteroaryl; or (ii) R^2 and R^3 together form alkylene, alkenylene or cycloalkylene; R^6 , R^7 and R^8 are each independently a monovalent group selected from hydrogen, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclalkyl, heterocyclalkenyl, heterocyclalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ;

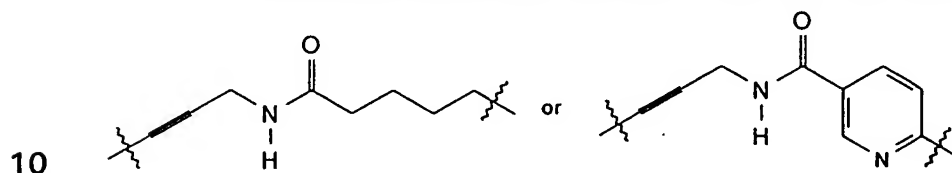
Y^2 is a salt of the oxyamino group, including but not limited to, mineral acids salts, such as but not limited to hydrochlorides and sulfates, and salts of organic acids, such as but not limited to acetates, lactates, malates, tartrates, citrates, ascorbates, succinates, butyrates, valerates and fumarates, or any amino protecting group, preferably an amino protecting group useful in the synthesis of oligonucleotides, more preferably monomethoxytrityl (MMT), dimethoxytrityl (DMT), 9-fluorenylmethoxycarbonyl (Fmoc), acetyl, trifluoroacetyl (TFA), benzoyl, or an oxime that is cleaved under mild acidic conditions (e.g., 100 mM acetate, pH 4.5-5.5) including, but not limited to, an oxime formed from a lower aliphatic aldehyde or ketone, preferably from acetone, cyclohexanone, propanal or 2-butanone;

where R^1 , R^2 , R^3 , R^4 , R^5 , R^6 , R^7 , R^8 , R^9 , R^{10} and Y^2 are unsubstituted or substituted with one or more substituents each independently selected from Z, as defined above.

In a preferred embodiment, the compounds for use in the methods provided herein are of formula (VII) where M has 1-50, preferably 1-25, more preferably 1-10 of the following groups, which can be combined in any order: arylene, heteroarylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$, $-C\equiv C-$, O, $S(A)_a$, $N(R^1)$, $N(COR^1)$ and $C(E)$; where a is 0, 1 or 2; A is O or NR^1 ; and E is S, O or NR^1 ; where R^1 is selected as above. In more preferred embodiments, M has 1-10 of the following groups, which can be combined in any order: $C(R^1)_2$, $-C\equiv C-$, $N(R^1)$ and $C(E)$; where E is S, O or

NR¹; where R¹ is selected as above. R¹ is preferably hydrogen; and E is preferably oxygen. In these more preferred embodiments, M includes 1-10, preferably 1-8, of the following groups, which can be combined in any order: CH₂, -C≡C-, NH and C(O).

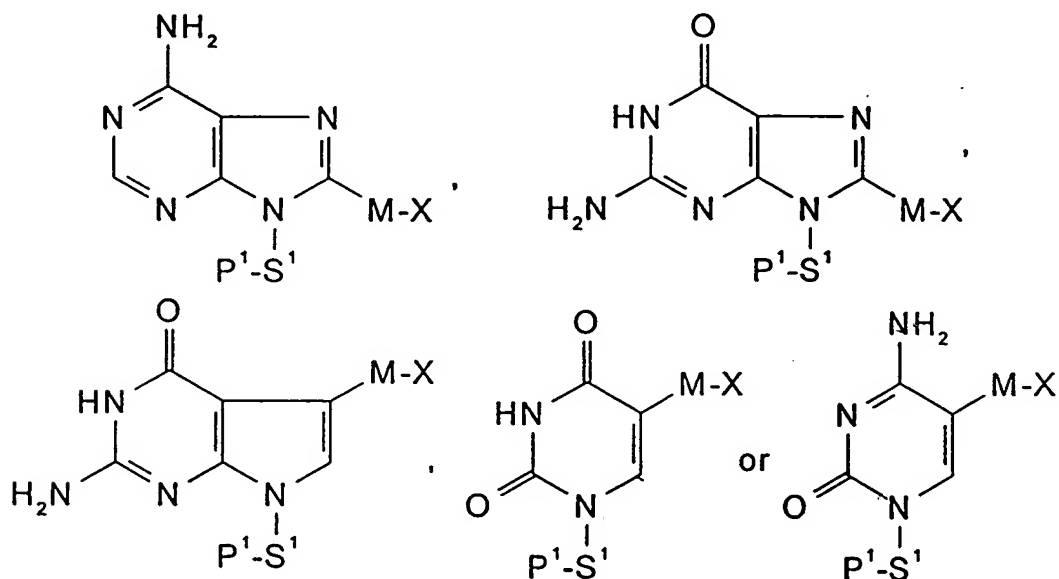
5 In more preferred embodiments, M has the formula:



4. Nucleobases B¹

The nucleobases B¹ of the monomers provided herein are well known to those of skill in the art, and include, but are not limited to, cytosines, uracils, adenines, guanines and thymines, and analogs thereof. including, but not limited to, deaza analogs, such as 7-deazaadenine, 7-deazaguanine and 7-deazahypoxanthine.

20 Thus, in preferred embodiments, the compounds of formula (I) have the formulae:



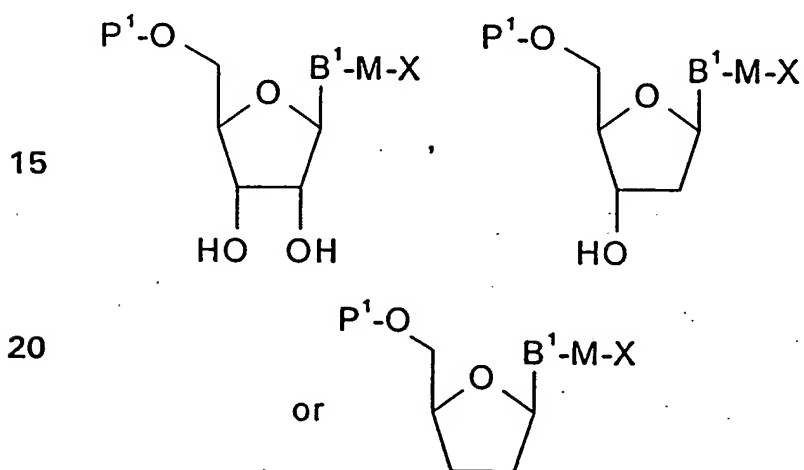
where P¹, S¹, M and X are selected as above.

5. Sugars S¹

The sugars S¹ of the monomers provided herein include, but are not limited to, riboses, deoxyriboses including 2'-deoxyriboses, and dideoxyriboses including 2',3'-dideoxyriboses. Other sugars for use

herein include arabinoses, including β -D-arabinofuranose, and the riboses 3'- β -D-ribofuranose, 3'-amino-2',3'-dideoxy- β -D-ribofuranose, 2',3'-dideoxy-3'-fluoro- β -D-ribofuranose and 2',3'-dideoxy-2',3'-didehydro- β -D-ribofuranose.

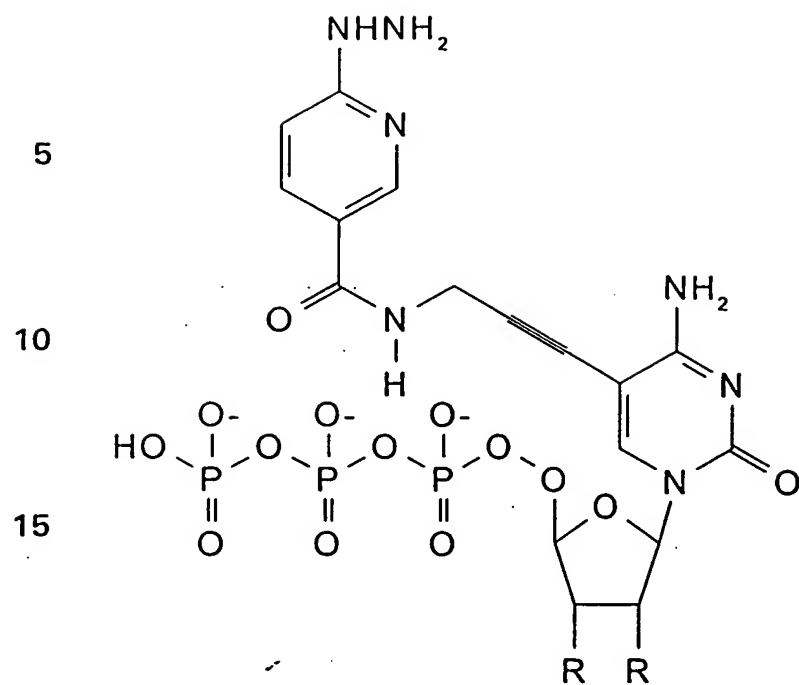
In preferred embodiments, the compounds of formula (I) have the formulae:



where P¹, B¹, M and X are selected as above.

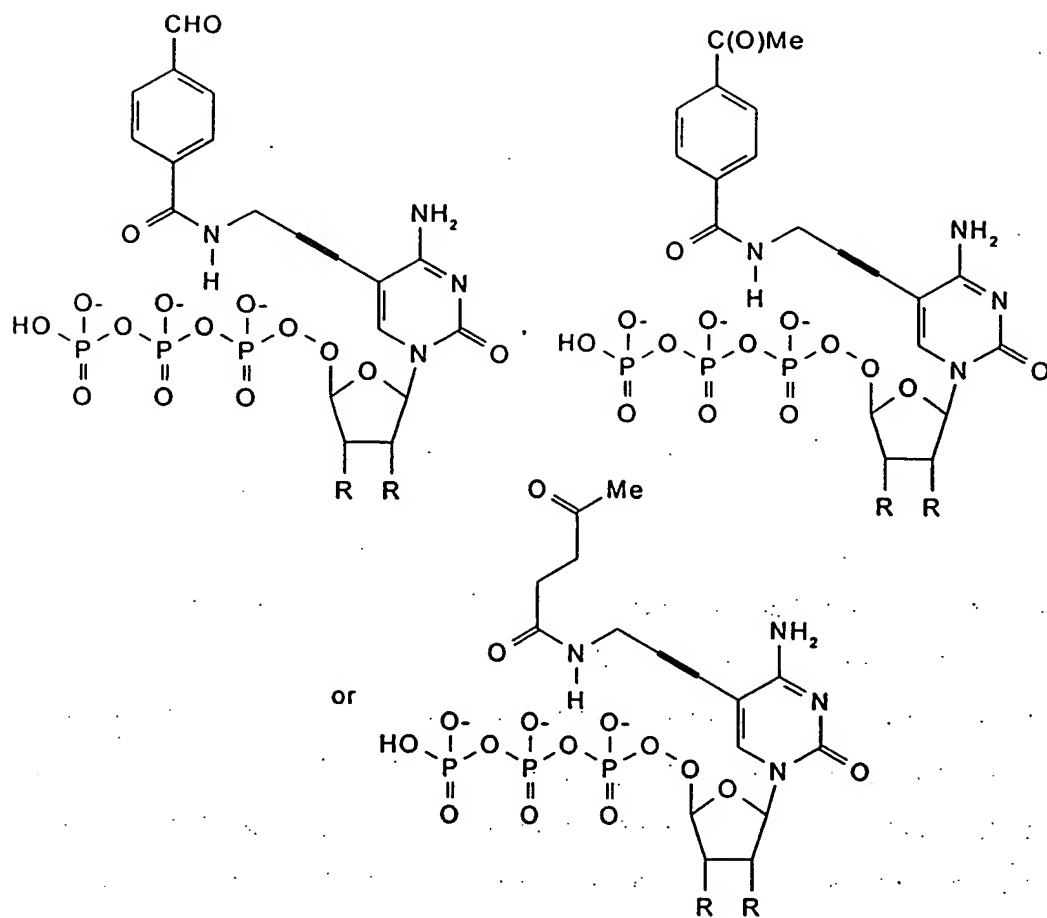
6. Preferred Monomers

The presently most preferred compounds of formula (IV) have the structure:



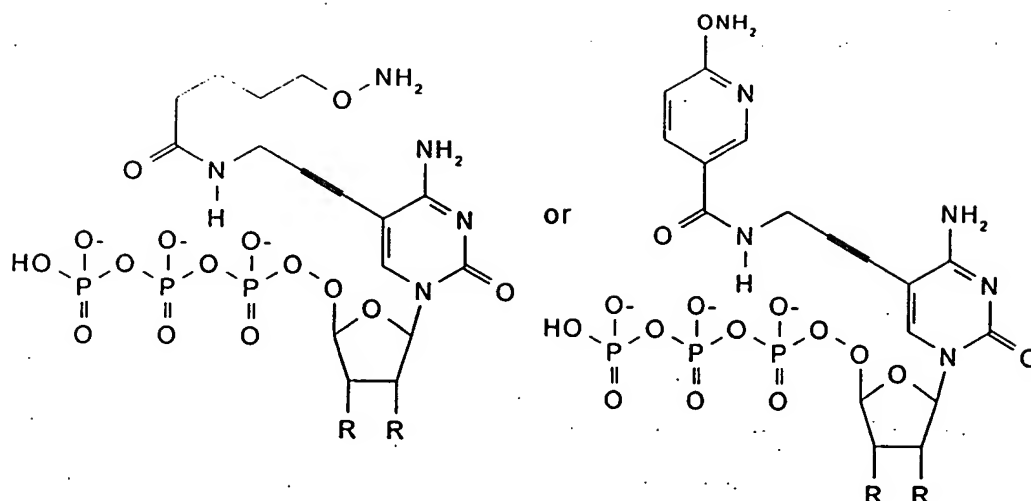
where each R is independently H or OH.

The presently most preferred compounds of formulae (VI) have the structures:



where each R is independently H or OH.

The presently most preferred compounds of formula (VII) have the structures:



where each R is independently selected from H and OH.

7. Methods of preparation of the Monomers

Methods of preparation of the monomers are provided. The methods involve the steps of: (i) derivatizing a carboxylic acid selected from an ω -carbonyl, an ω -protected hydrazino, and an ω -protected oxyamino substituted carboxylic acid as the corresponding active ester, including, but not limited to, aryl, heteroaryl, succinimidyl, maleimidyl, phthalimidyl and naphthimidyl, preferably succinimidyl, esters by reacting the carboxylic acid with the appropriate aryl, heteroaryl, succinimidyl, maleimidyl, phthalimidyl or naphthimidyl alcohol, preferably N-hydroxysuccinimide, under dehydrating conditions, including, but not limited to, in the presence of a carbodiimide, such as dicyclohexylcarbodiimide; and (ii) reacting the resulting active ester with an amino substituted nucleoside triphosphate, including but not limited to a 5-amino substituted cytidine triphosphate such as 5-1, as shown in FIGURE 5. Preferred ω -carbonyl carboxylic acids include 4-formylbenzoic acid, levulinic acid and 4-carboxyacetophenone. Preferred ω -protected hydrazino carboxylic acids include 6-(N'-dimethoxytrityl)hydrazinonicotinic acid. Preferred ω -protected oxyamino carboxylic acids include 6-(N-dimethoxytrityl)oxyaminonicotinic acid.

B. Modified Oligonucleotides, Immobilized Oligonucleotides and Oligonucleotide Conjugates

1. Modified Oligonucleotides

Provided herein are modified oligonucleotides that possess a

5 hydrazino group (see, FIGURE 2), an oxyamino group, or a carbonyl group. These modified oligonucleotides are synthesized directly during enzymatic synthesis without the need for post synthetic modification to produce the hydrazino, oxyamino, or carbonyl group. The modified oligonucleotide first components can be immobilized or conjugated

10 through the formation of covalent hydrazone or oxime bonds. Such hydrazone or oxime bonds can be formed by reaction of the hydrazino, oxyamino or carbonyl modified oligonucleotide first components with appropriately derivatized solid surfaces, as described herein; or second components, including, but not limited to, biopolymers as defined herein,

15 polymers including, but not limited to, polyamines, polyamides, polyethers and polyethylene glycols, and other compounds of interest herein for use in assays, kits, diagnostic arrays, and the like, including, but not limited to, intercalators, vitamins, reporter molecules, cholesterol, phospholipids, biotin, phenazine, phenanthridine,

20 anthraquinone, acridine, fluoresceins, rhodamines, coumarins, dyes, antibodies, haptens, antigens, enzymes, and detection reagents including, but not limited to, fluorophores, metals including, but not limited to, gold, metal chelates, chromophores, fluorophore precursors and chromophore precursors, that are modified to possess a hydrazino, oxyamino or

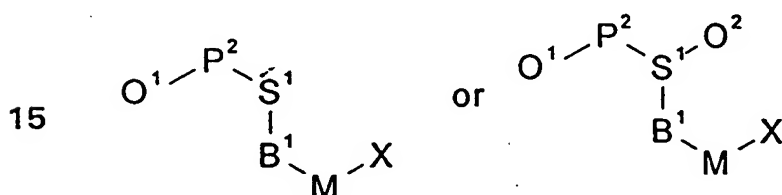
25 carbonyl group that is complementary to the carbonyl, oxyamino or hydrazino group of the oligonucleotide of formula (II) for formation of hydrazone or oxime linkage. Dendrimeric compounds that possess a plurality of detectable groups, including, but not limited to, reporter molecules, fluorophores, chromophores, fluorophore precursors and

30 chromophore precursors, are also contemplated herein as second components. Fluorophore precursors and chromophore precursors are

compounds that react with the hydrazino, oxyamino or carbonyl group of the modified oligonucleotide to form a fluorogenic or chromogenic group for analysis. Fluorophore precursors and chromophore precursors are well known to those of skill in the art, and include, but are not limited to, 3-(4-carboxybenzoyl)quinoline-2-carboxaldehyde (CBQCA) and 3-(2-furoyl)quinoline-2-carboxaldehyde (FQ), sold by Molecular Probes, Inc. (Eugene, Oregon).

Importantly, the reaction to form the immobilized oligonucleotide or oligonucleotide conjugate can be performed under aqueous conditions without the need for additional reagents, such as a reducing agent.

The modified oligonucleotides provided herein have formula (II):



20 or a derivative thereof, where O^1 and O^2 are each independently an oligonucleotide or analog thereof, such as a protein nucleic acid (PNA); P^2 is a phosphorous linking group resulting from the coupling of a compound of formula (I) with the oligonucleotide or analog thereof, preferably a phosphodiester group; and S^1 , B^1 , M and X are selected as above for the hydrazino, oxyamino, and carbonyl monomers.

In certain embodiments, the oligonucleotide analogs of formula (II) are selected with the proviso that if X is CHO , then M is not undecylenecarbonylaminomethylene ($n\text{-C}_{11}\text{H}_{22}\text{C}(\text{O})\text{NHCH}_2$) or C_{1-20} alkylene. In other embodiments, the oligonucleotide analogs are of formula (II) where X is not an oxyamino or protected oxyamino group. In further embodiments, the oligonucleotide analogs are of formula (II) where X is not a hydrazide ($-\text{C}(\text{O})\text{-NHNH}_2$) group. In particular embodiments, the oligonucleotide analogs of formula (II) are selected with the proviso that if X is $-\text{NHNH}_2$, then M is not $-\text{CH}(\text{CH}_2\text{OH})(\text{CH}_2)_6\text{NHCO-5-}$

pyrid-2-ylene. In other embodiments, the oligonucleotide analogs of formula (II) are selected with the proviso that P^2 is not a phosphoramidate group.

Thus, in more preferred embodiments, the modified

5 oligonucleotides have formulae:

10

15

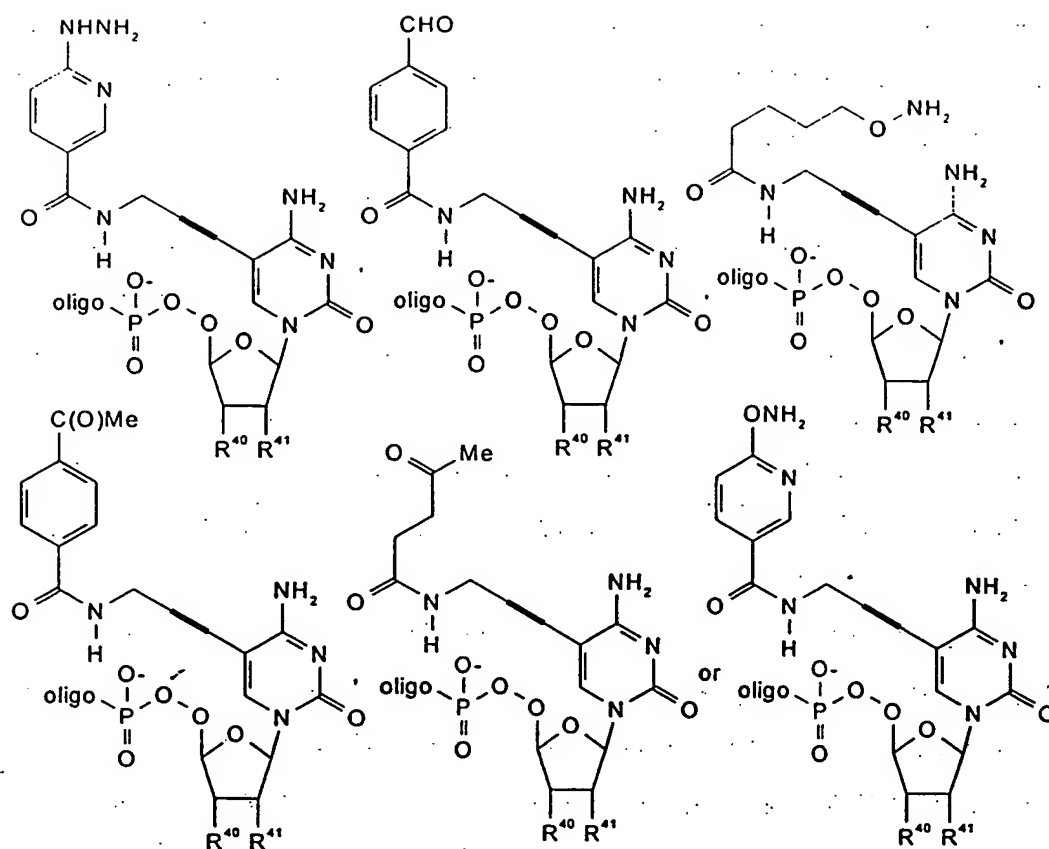
20

25

30

35

40



where R^{40} is selected from an oligonucleotide, OH and H; and R^{41} is selected from OH and H.

Methods of synthesis of the modified oligonucleotides are also provided. The methods involve the step of contacting a composition having at least one nucleoside triphosphate provided herein with an enzyme for oligonucleotide synthesis under conditions where oligonucleotide synthesis occurs. Preferred enzymes include reverse transcriptases, including, but not limited to, AMV reverse transcriptase, MMLV reverse transcriptase and superscript reverse transcriptase, and polymerases, including, but not limited to, Taq polymerase, DNA polymerase, Klenow fragment and T4 DNA polymerase.

The modified oligonucleotides provided herein are also useful in studying how a normal cell differs for its diseased state. Current gene expression studies require the determination of the expression of any new genes as well as the difference in expression of genes in the two cell

states. To this end, RNA from the two cell populations are isolated and cDNA libraries are prepared using reverse transcriptase enzymes and fluorescently labeled triphosphates such as Cy3 and Cy5 triphosphates. It is recognized that the stability and kinetics of incorporation of the

5 labeled triphosphates are not identical. The modified triphosphates provided herein may be used in both applications. This is advantageous over current methodologies, e.g., amino substituted nucleoside triphosphates (Clontech) labeled with succinimidyl fluorophores, due to the inherent instability of the succinimidyl esters and the low

10 concentration of amino groups incorporated in the cDNA product.

2. Immobilized Oligonucleotides

Oligonucleotides immobilized through a hydrazone or oxime linkage to a solid surface are provided. The oligonucleotides are of formula (II), and are modified with a hydrazino, oxyamino, or carbonyl moiety.

15 Reaction of the modified oligonucleotide with an appropriately functionalized solid surface provides an immobilized oligonucleotide for use in preparation of microarrays, diagnostic probe assays, DNA amplification by solid phase polymerase chain reactions (PCR), molecular computing (see, e.g., Adleman (1994) Science 266:1021-1024; Kari

20 (1997) Mathematical Intelligencer 19:9-22; Frutos et al. (1997) Nucleic Acids Res. 25:4748; Smith et al. (1998) J. Comp. Biol. 5:255; Liu et al. (1998) J. Comp. Biol. 5:267; Frutos et al. (1998) J. Am. Chem. Soc. 120:10277; Wang et al. (1999) Biosystems 52:189-191; Liu et al. (1999) Biosystems 52:25-33; Liu et al. (2000) Nature 403:175-179;

25 European Patent Application Publication No. EP 0 772 135; Reed et al. (June 2000) Scientific American:86-93), molecular addressing (Niemeyer et al. (1994) Nucl. Acids Res. 22(25):5530-5539), DNA sequencing by mass spectrometry (see, e.g., U.S. Patent Nos. 6,074,823 and 5,547,835) and in studying the molecular electronics of DNA (see, e.g.,

30 U.S. Patent Nos. 6,071,699, 6,066,448, 5,952,172 and 5,824,473).

Importantly, the reaction to form the immobilized oligonucleotide can be performed under aqueous conditions without the need for additional reagents, such as a reducing reagent.

The modified solid surfaces are prepared on a variety of supports,
5 as described herein. Appropriately modified solid surfaces can be prepared by the methods provided herein and possess carbonyl, oxyamino, or hydrazino groups complementary to the modified oligonucleotide. Reaction of the modified oligonucleotide with the modified solid surface results in covalent attachment of the
10 oligonucleotide to the solid surface through a hydrazone or oxime linkage.

In embodiments where X of the modified oligonucleotide is a hydrazino or oxyamino group, the solid surface can be modified to possess an epoxide, α -bromocarbonyl, maleimide, maleic anhydride, isothiocyanate or isocyanate group. Such solid surfaces can be prepared
15 by methods provided herein or other methods well known to those of skill in the art. For example, reaction of pentafluorophenyl 4-isothiocyanatobenzoate with an amino or hydroxy solid surface results in formation of an isothiocyanato modified solid surface. Some of these surfaces are commercially available from, e.g., Pierce (Rockford, Illinois), SINTEF
20 Applied Chemistry (Trondheim, Norway), Rapp Polymere GmbH (Tubingen, Germany), and Dyno Particles AS (Trondheim, Norway). Reaction of the hydrazino or oxyamino group of the modified oligonucleotide with the epoxide, α -bromocarbonyl, maleimide, maleic anhydride, isothiocyanate or isocyanate group of the solid surface results
25 in covalent attachment of the oligonucleotide to the solid surface.

In certain embodiments, particularly where X is an oxyamino group, the immobilized oligonucleotides are selected such that the solid surface is not modified with an aldehyde or epoxide group.

3. Oligonucleotide Conjugates

Oligonucleotide first components conjugated to second components, including, but not limited to, biopolymers as defined herein, polymers including, but not limited to, polyamines, polyamides, polyethers and polyethylene glycols, and other compounds of interest herein for use in assays, kits, diagnostic arrays, and the like, including, but not limited to, intercalators, vitamins, reporter molecules, cholesterol, phospholipids, biotin, phenazine, phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, dyes, antibodies, haptens, antigens, enzymes, and detection reagents including, but not limited to, fluorophores, metals including, but not limited to, gold, metal chelates, chromophores, fluorophore precursors and chromophore precursors, that are modified to possess a hydrazino, oxyamino or carbonyl group that is complementary to the carbonyl, oxyamino or hydrazino group of the modified oligonucleotide of formula (II) for formation of hydrazone or oxime linkages are provided. Dendrimeric compounds that possess a plurality of detectable groups, including, but not limited to, reporter molecules, fluorophores, chromophores, fluorophore precursors and chromophore precursors, are also contemplated herein as second components. Fluorophore precursors and chromophore precursors are compounds that react with the hydrazino, oxyamino or carbonyl group of the modified oligonucleotide to form a fluorogenic or chromogenic group for analysis.

Reaction of a modified oligonucleotide of formula (II) with an appropriately modified second component provides a hydrazone or oxime linked conjugate. Thus, in one embodiment, a hydrazino modified oligonucleotide of formula (II) reacts with a carbonyl containing second component, or a second component modified to possess a carbonyl group, such as an aldehyde or a ketone, to afford the conjugate provided herein. In another embodiment, a carbonyl modified oligonucleotide of formula (II) is reacted with a second component containing a hydrazino

group, or modified to contain a hydrazino group, to afford the conjugate provided herein. In other embodiments, the modified oligonucleotide possesses an oxyamino or carbonyl group, and the second component possesses a complementary carbonyl or oxyamino group, or is modified
 5 to possess a complementary carbonyl or oxyamino group, to afford an oxime linked conjugate.

Importantly, the reaction to form the oligonucleotide conjugate can be performed under aqueous conditions without the need for additional reagents, such as a reducing agent.

10 C. Hydrazino, Oxyamino and Carbonyl Modified Surfaces

Provided herein are reagents and methods for modification of solid surfaces and second components for use in the methods provided herein. In particular, the modified solid surfaces and second components are useful for the immobilization of the oligonucleotide first components
 15 provided herein, and in the conjugation of such first components, respectively.

It is to be understood that the reagents described below are interchangeable in certain applications. For example, a dihydrazine reagent for modification of oligosaccharides can also be used in the
 20 modification of solid surfaces to provide a hydrazino solid surface. Many of the other reagents provided below can be used in the preparation of either modified solid surfaces or in the modification of second components for conjugation with the modified oligonucleotides provided herein.

25 1. Modified Surfaces

Hydrazino, oxyamino and carbonyl modified surfaces are provided. The surfaces provided herein include, but are not limited to, glasses including glasses including controlled pore glass, plastics including polyethylenes, polycarbonates, polypropylenes, polyamides such as
 30 Nylon® and polyacrylamides, and polyvinylidenedifluorides, latexes, rubbers, celluloses including nitrocelluloses and diazotized celluloses,

- latexes, metals including platinum, nickel, zinc, tin, palladium, steel, gold, silver, aluminum, silicon and copper, silicas, agaroses including Sephadex, dextrans including Sepharose, natural sponges, polystyrenes including polystyrene crosslinked with divinylbenzene, polystyrenes
- 5 radiation grafted onto perfluorinated polymers such as Teflon[®] (see, e.g., Maeji et al. (1994) Reactive Polymers 22:203-212; and Berg et al. (1989) J. Am. Chem. Soc. 111:8024-8026), oxides of metals and metalloids such as Pt-PtO, Si-SiO, Au-AuO, TiO₂, Cu-CuO, and the like, and compound semiconductors, such as lithium niobate, gallium arsenide and
 - 10 indium-phosphide. Other solid surfaces for use herein are well known to those of skill in the art. See, e.g., U.S. Patent Nos. 4,507,230, 4,006,117, 5,389,449, 5,556,752, 4,683,202 and 5,744,305; International Patent Application Publication Nos. WO 00/04382; WO 00/04390 and WO 00/04389; Merrifield (1964) Biochemistry
 - 15 3:1385-1390; Berg et al. (1990) in Innovation Perspect. Solid Phase Synth. Collect. Pap., Int. Symp., 1st, Epton, Roger (Ed), pp. 453-459; Berg et al. (1989) in Pept., Proc. Eur. Pept. Symp., 20th, Jung, G. et al. (Eds), pp. 196-198; Berg et al. (1989) J. Am. Chem. Soc. 111:8024-8026; Kent et al. (1979) Isr. J. Chem. 17:243-247; Kent et al.
 - 20 (1978) J. Org. Chem. 43:2845-2852; Mitchell et al. (1976) Tetrahedron Lett. 42:3795-3798; and Hermanson et al. (1992) Immobilized Affinity Ligand Techniques, Academic Press, Inc., San Diego.

- Synthetic solid surfaces include those made from polymers and co-polymers such as polyvinylalcohols, acrylates and acrylic acids such as
- 25 polyethylene-co-acrylic acid, polyethylene-co-methacrylic acid, polyethylene-co-ethylacrylate, polyethylene-co-methyl acrylate, polypropylene-co-acrylic acid, polypropylene-co-methyl-acrylic acid, polypropylene-co-ethyl-acrylate, polypropylene-co-methyl acrylate, polyethylene-co-vinyl acetate, polypropylene-co-vinyl acetate, and those containing acid anhydride
 - 30 groups such as polyethylene-co-maleic anhydride, polypropylene-co-maleic anhydride and the like. Liposomes have also been used as solid

surfaces for affinity purifications (see, e.g., Powell et al. (1989) Biotechnol. Bioeng. 33:173).

For example, U.S. Patent No. 5,403,750, describes the preparation of polyurethane-based polymers. U.S. Pat. No. 4,241,537 describes a
5 plant growth medium containing a hydrophilic polyurethane gel composition prepared from chain-extended polyols; random copolymerization is preferred with up to 50% propylene oxide units so that the prepolymer will be a liquid at room temperature. U.S. Pat. No. 3,939,123 describes lightly crosslinked polyurethane polymers of isocyanate terminated
10 prepolymers comprised of poly(ethyleneoxy) glycols with up to 35% of a poly(propyleneoxy) glycol or a poly(butyleneoxy) glycol. In producing these polymers, an organic polyamine is used as a crosslinking agent. Other solid surfaces and preparation thereof are described in U.S. Patent Nos. 4,177,038, 4,175,183, 4,439,585, 4,485,227, 4,569,981,
15 5,092,992, 5,334,640, and 5,328,603.

U.S. Patent No. 4,162,355 describes a polymer suitable for use in affinity chromatography, which is a polymer of an aminimide and a vinyl
compound having at least one pendant halo-methyl group. An amine
ligand, which affords sites for binding in affinity chromatography is
20 coupled to the polymer by reaction with a portion of the pendant halo-methyl groups and the remainder of the pendant halo-methyl groups are reacted with an amine containing a pendant hydrophilic group. A method of coating a substrate with this polymer is also described. An exemplary aminimide is 1,1-dimethyl-1-(2-hydroxyoctyl)amine
25 methacrylimide and vinyl compound is a chloromethyl styrene.

U.S. Patent No. 4,171,412 describes specific matrices based on hydrophilic polymeric gels, preferably of a macroporous character, which carry covalently bonded D-aminoacids or peptides which contain
D-aminoacid units. The basic carrier is prepared by copolymerization of
30 hydroxyalkyl esters or hydroxyalkylamides of acrylic and methacrylic acid with crosslinking acrylate or methacrylate comonomers are modified by

the reaction with diamines, aminoacids or dicarboxylic acids and the resulting carboxyterminal or aminoterminal groups are condensed with D-analogs of aminoacids or peptides. The peptide containing D-aminoacids also can be synthesized stepwise on the surface of the
5 carrier.

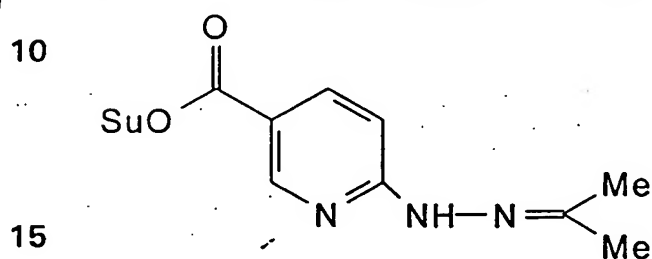
U.S. Patent No. 4,178,439 describes a cationic ion exchanger and a method for preparation thereof. U.S. Patent No. 4,180,524 describes chemical syntheses on a silica support.

The solid supports can be identifiable, such as by with electronic
10 tags or bar codes (see, *e.g.*, U.S. Patent No. 6,025,129; U.S. Patent No. 6,017,496; U.S. Patent No. 5,972,639; U.S. Patent No. 5,961,923; U.S. Patent No. 5,925,562; U.S. Patent No. 5,874,214; U.S. Patent No. 5,751,629; U.S. Patent No. 5,741,462), or chemical tags (see, U.S. Patent No. 5,432,018; U.S. Patent No. 5,547,839), addressable arrays
15 and other such addressing methods. Supports with electronic tags include supports with radio-frequency tags, such as IRORI MICROKANS® and MICROTUBES® microreactors (see, U.S. Patent No. 6,025,129; U.S. Patent No. 6,017,496; U.S. Patent No. 5,972,639; U.S. Patent No. 5,961,923; U.S. Patent No. 5,925,562; U.S. Patent No. 5,874,214; U.S.
20 Patent No. 5,751,629; U.S. Patent No. 5,741,462; International PCT application No. WO98/31732; International PCT application No. WO98/15825; and, see, also U.S. Patent No. 6,087,186).

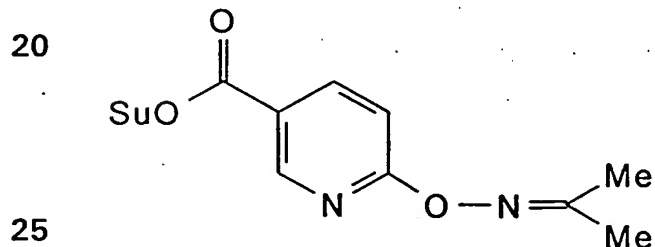
The solid surfaces can take the form of beads, capillaries, flat supports such as glass fiber filters, multiwell plates, thin films,
25 membranes, wafers, combs, pins, beads in pits of flat surfaces such as wafers (*e.g.*, silicon wafers), with or without filter plates, pellets, disks, continuous surfaces such as a microtiter dish or well, hollow fibers, needles, solid fibers, slides, chips, sheets, containers or meshes. The solid surfaces can be formed in arrays for combinatorial synthesis or
30 analysis, or for formation of a combinatorial library. Other supports for use herein include paper plus coated with a suitable matrix

The solid surface is modified to possess a hydrazino, oxyamino, or carbonyl group that is complementary to the hydrazino, oxyamino or carbonyl moiety of the oligonucleotide of formula (II) for formation of a hydrazone or oxime linkage.

5 The hydrazino or oxyamino modified surfaces are prepared from commercially available solid supports or known supports such as those described above. Reaction of the support with, e.g., 5-succinimidylloxycarbonylpyrid-2-yl acetone hydrazone having the formula:

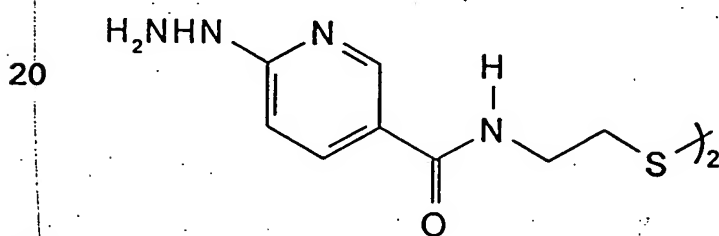


or the corresponding oxyamino compound having the formula:



forms a hydrazone or oxime modified solid surface. Alternatively, the hydrazone or oxime succinimidyl esters shown above are reacted with an amino-substituted silane, such as 3-aminopropyltriethoxysilane, to form a silyl hydrazone or oxime. Reaction of the silyl group with a solid surface forms a hydrazone or oxime modified solid surface. Hydrolysis of the hydrazone or oxime group under mild acidic conditions (e.g., 0.1 M acetate, pH 4.7) affords the hydrazino or oxyamino modified surfaces provided herein. See, co-pending, co-owned U.S. provisional application Serial No. 60/191,186, the disclosure of which is incorporated herein by reference in its entirety.

In one embodiment, hydrazino modified gold surfaces are prepared by reaction with a dihydrazino-dithiol compound of formula $X-M-S-S-M-X$, where each S is sulfur, each X is independently a protected or unprotected hydrazino group, and each M is independently selected as
 5 above. In preferred embodiments, the compound is symmetrical and has the formula $(X-M-S)_2$. In these embodiments, M preferably has any combination, preferably 1 to 100, more preferably 1 to 50, most preferably 1 to 10, of the following groups: arylene, heteroarylene, $C(E)$, $N(R^1)$ and $C(R^1)_2$, where E and R^1 are selected as above, which groups
 10 can be combined in any order. R^1 is preferably hydrogen and E is preferably oxygen. In more preferred embodiments, M has 1 to 10 of the following groups: pyridylene, NH, $C(O)$ and CH_2 , which groups can be combined in any order. M is most preferably derived from cysteamine. Briefly, reaction of cysteamine dimer (*i.e.*, $(H_2NCH_2CH_2S)_2$) with
 15 succinimidyl 2-BOC-hydrazino-5-pyridylcarboxylate, followed by deprotection with, *e.g.*, dilute hydrochloric acid, provides the dihydrazino-dithiol:



Reaction of this dithiol with a gold surface results in cleavage of the dithiol bond to form a hydrazino modified gold surface (*i.e.*, $Au-S-CH_2CH_2-NHC(O)-pyr-NHNH_2$).

30 Carbonyl modified surfaces are prepared from commercially available solid supports such as those described above. Reaction of the support with, *e.g.*, succinimidyl 4-formylbenzoate or 4'-succinimidyl-oxycarbonylacetophenone forms a carbonyl, particularly an aldehyde or methyl ketone modified solid surface. Alternatively, succinimidyl 4-formyl

benzoate or 4'-succinimidylloxycarbonylacetophenone is reacted with an amino-substituted silane, such as 3-aminopropyltriethoxysilane, to form a silyl aldehyde or ketone. Reaction of the silyl group with a solid surface forms an aldehyde or ketone modified solid surface.

5 2. Modified Second Components

Hydrazino, oxyamino, and carbonyl modified second components can be prepared by methods provided herein, or by methods well known to those of skill in the art.

a. Modified Oligosaccharides and Proteins

10 For example, an oligosaccharide can be reacted with cyanogen bromide or 1-cyano-4-dimethylaminopyridinium bromide (CDAP) to afford a cyano ether of formula $R-O-CN$. Reaction of this cyano ether with a hydrazino, oxyamino or carbonyl substituted amine, provided herein, affords a hydrazino, oxyamino or carbonyl substituted oligosaccharide, 15 where the hydrazino, oxyamino or carbonyl group is linked to the oligosaccharide through a group of formula $-O-C(=NH)-NH-$.

In one embodiment, the hydrazino, oxyamino or carbonyl substituted amine provided herein for oligosaccharide modification has the formula:

20 H_2N-M-X ,

or a derivative thereof, where M and X are as defined above. In certain embodiments, the compound of formula H_2N-M-X is selected where M is not C_{1-12} alkyl.

In the above embodiments, the modified oligosaccharide has the 25 formula:

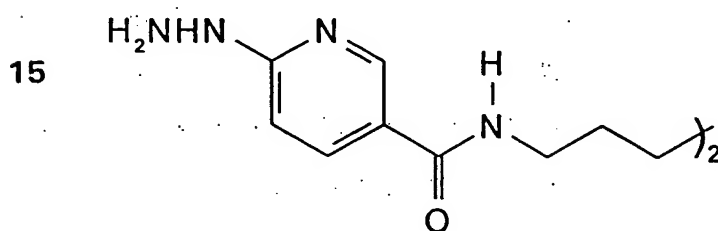
$O^2-O-C(=NH)-NH-M-X$,

where O^2 is an oligosaccharide; and M and X are as defined above.

Alternatively, oligosaccharides are derivatized as their hydrazino analogs by reaction with CDAP and a dihydrazino compound. The 30 dihydrazino compounds have the formula:

$X-M-X$,

or a derivative thereof, where each X is independently a hydrazino group and M is selected as above. In preferred embodiments, M has any combination, preferably 1 to 100, more preferably 1 to 50, most preferably 1 to 10, of the following groups: arylene, heteroarylene, C(R¹)₂, N(R¹) and C(E), where R¹ and E are selected as above, which groups can be combined in any order. In preferred embodiments, R¹ is hydrogen and E is oxygen. In more preferred embodiments, M has 1 to 10 of the following groups: heteroarylene, CH₂, NH and C(O), which groups can be combined in any order. M most preferably is derived, in these embodiments, from reaction of 1,6-hexanediamine with two equivalents of succinimidyl N'-BOC-6-hydrazinonicotinate. Thus, the reagent of formula X-M-X most preferably has formula:



The resulting hydrazino modified oligosaccharide is then conjugated to a carbonyl modified oligonucleotide provided herein. Similarly, di(oxyamino) compounds of formula X-M-X, or a derivative thereof, where X is an oxyamino group and M is selected as above, can be used to generate oxyamino modified oligosaccharides for conjugation with carbonyl modified oligonucleotides.

Polysaccharides or glycoproteins can be oxidized with sodium periodate to yield aldehyde groups directly. These aldehyde groups can be directly labeled with hydrazine modified oligonucleotides as described herein. Lysine moieties on proteins can be modified to incorporate aldehydes using succinimidyl 4-formylbenzoate (Pierce Chemical Co.). These aldehyde containing proteins can be directly conjugated to hydrazine modified oligonucleotides.

Hydrazine groups can also be incorporated on insoluble polysaccharides such as cellulose or agarose using 1-cyano-4-dimethylaminopyridine bromide (CDAP) as described by Lees et al. ((1996) Vaccine 14:190 and (2000) Vaccine 18:1273) using a dihydrazine molecule (see, e.g., Example 16).

Modified proteins possessing hydrazino or oxyamino groups can be prepared by the methods described in U.S. Patent Nos. 5,206,370, 5,420,285 and 5,753,520, and European Patent Specification EP 0 384 769 B1, the disclosures of which are incorporated herein by reference in their entirety. These hydrazine modified proteins can be directly conjugated to aldehyde modified oligonucleotides prepared as described herein. Carbonyl modified proteins are prepared by reaction of an aldehyde or ketone substituted succinimidyl ester with nucleophilic groups of the protein, including, but not limited to, the amino side chain of lysine. In certain embodiments, the aldehyde or ketone substituted N-hydroxysuccinic ester is, for example, succinimidyl 4-formyl- or 4-acetylbenzoate. In other embodiments, the protein is reacted with succinimidyl levulinate (succinimidyl 4-oxopentanoate) to provide a ketone substituted protein.

Methods of synthesis of hydrazino, oxyamino, and carbonyl modified oligonucleotides are provided herein.

In other embodiments herein, particularly when the modified oligonucleotide possesses a hydrazino or oxyamino group, the second component can possess an epoxide, α -bromocarbonyl, maleimide, maleic anhydride, isothiocyanate or isocyanate group. Such second components can be prepared by methods provided herein or other methods well known to those of skill in the art. For example, reaction of pentafluorophenyl 4-isothiocyanatobenzoate with an amino or hydroxy group of a second component results in formation of an isothiocyanato modified solid surface. Reaction of the hydrazino or oxyamino group of the modified oligonucleotide with the epoxide, α -bromocarbonyl,

maleimide, maleic anhydride, isothiocyanate or isocyanate group of the second component results in covalent attachment of the oligonucleotide to the second component. In these embodiments, the conjugation reaction with a hydrazino modified oligonucleotide can be conducted at a lower pH than for the corresponding amino modified oligonucleotide.

b. Modified Indicators

Modified indicators can be prepared according to the methods provided herein. The modified indicators are conjugated with the modified oligonucleotides of formula (II) to form a hydrazino or oxime linkage. Preferred indicators include, but are not limited to, fluoresceins, rhodamines and cyanine dyes. The modified indicators possess a hydrazino, oxyamino or carbonyl group that is complementary to the carbonyl, oxyamino or hydrazino group of the monomers of formula (I) or the modified oligonucleotides of formula (II).

In one embodiment, the modified indicators have the formula:



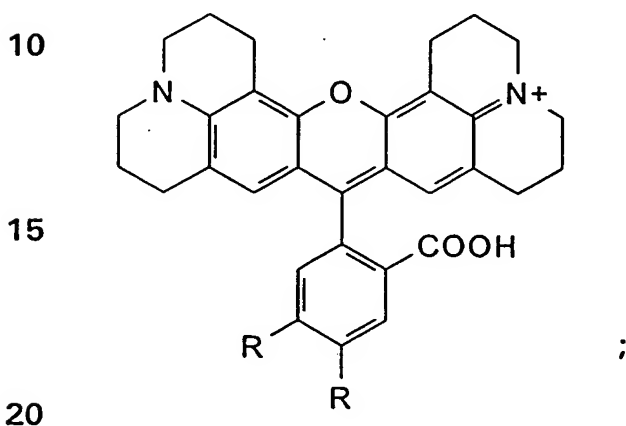
or a derivative thereof, where F^* is a fluorescein, a rhodamine or a cyanine dye; X is a protected or unprotected hydrazino group, a protected or unprotected oxyamino group or a carbonyl group; and

M is a divalent group having any combination of the following groups, which can be combined in any order: arylene, heteroarylene, cycloalkylene, $C(R^1)_2$, $-C(R^1)=C(R^1)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^1)$, $P(D)_b(ER^1)$, $N(R^1)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^1 ; D is S or O; and E is S, O or NR^1 ; each R^1 is a monovalent group independently selected from hydrogen and M^1-R^4 ; each M^1 is a divalent group independently having any combination of the following groups, which groups can be combined in any order: a direct link, arylene, heteroarylene, cycloalkylene, $C(R^5)_2$, $-C(R^5)=C(R^5)-$, $>C=C(R^2)(R^3)$, $>C(R^2)(R^3)$, $-C\equiv C-$, O, $S(A)_a$, $P(D)_b(R^5)$, $P(D)_b(ER^5)$, $N(R^5)$, $N(COR^5)$, $>N^+(R^2)(R^3)$ and C(E); where a is 0, 1 or 2; b is 0, 1, 2 or 3; A is O or NR^5 ; D is S or O; and E is S, O or NR^5 ; R^4 and R^5 are each

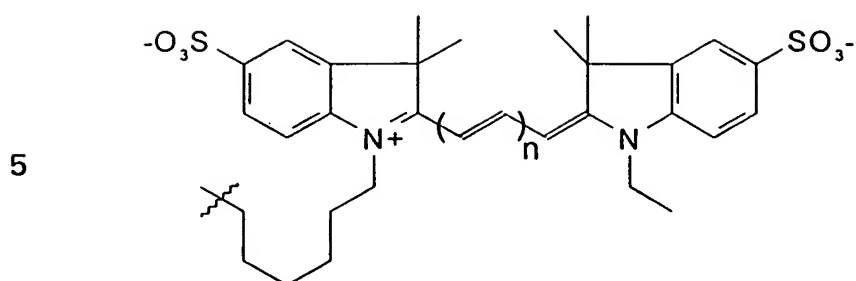
- independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, $\text{SiR}^6\text{R}^7\text{R}^8$, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; R^9 and R^{10} are each independently selected from hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, aralkyl, heteroaryl, heteroaralkyl and heterocyclyl; R^2 and R^3 are selected from (i) or (ii) as follows: (i) R^2 and R^3 are independently selected from among hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl and heteroaryl; or (ii) R^2 and R^3 together form alkylene, alkenylene or cycloalkylene; R^6 , R^7 and R^8 are each independently a monovalent group selected from hydrogen, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy and NR^9R^{10} ; and
- R^1 , R^2 , R^3 , R^4 , R^5 , R^6 , R^7 , R^8 , R^9 and R^{10} are unsubstituted or substituted with one or more substituents each independently selected from Z, wherein Z is selected from alkyl, alkenyl, alkynyl, aryl, cycloalkyl, cycloalkenyl, hydroxy, $\text{S}(\text{O})_h\text{R}^{20}$, $\text{NR}^{20}\text{R}^{21}$, COOR^{20} , COR^{20} , $\text{CONR}^{20}\text{R}^{21}$, $\text{OC}(\text{O})\text{NR}^{20}\text{R}^{21}$, $\text{N}(\text{R}^{20})\text{C}(\text{O})\text{R}^{21}$, alkoxy, aryloxy, heteroaryl, heterocyclyl, heteroaryloxy, heterocyclyoxy, aralkyl, aralkenyl, aralkynyl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, aralkoxy, heteroaralkoxy, alkoxycarbonyl, carbamoyl, thiocarbamoyl, alkoxycarbonyl, carboxyaryl, halo, pseudohalo, haloalkyl and carboxamido; h is 0, 1 or 2; and R^{20} and R^{21} are each independently selected from among hydrogen, halo, pseudohalo, cyano, azido, nitro, trialkylsilyl, dialkylarylsilyl, alkyl diarylsilyl, triarylsilyl, alkyl, alkenyl, alkynyl, haloalkyl, haloalkoxy, aryl, aralkyl, aralkenyl, aralkynyl, heteroaryl, heteroaralkyl, heteroaralkenyl, heteroaralkynyl, heterocyclyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, hydroxy, alkoxy, aryloxy, aralkoxy, heteroaralkoxy,

amino, amido, alkylamino, dialkylamino, alkylarylamino, diarylamino and arylamino.

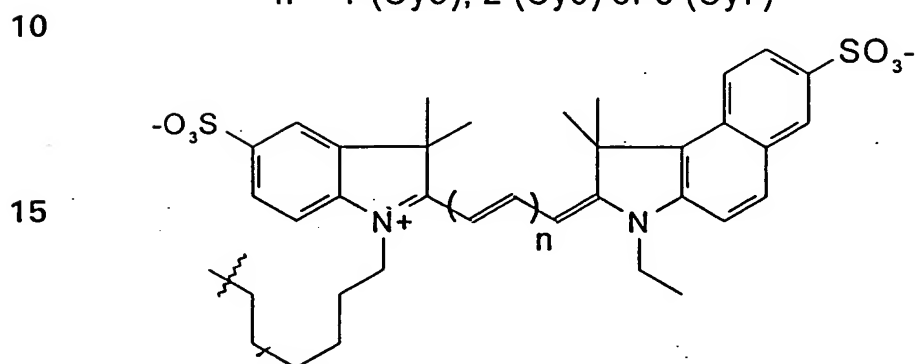
In preferred embodiments herein, the modified indicators are those where the fluorescein is 5-aminofluorescein, 6-aminofluorescein, fluorescein 5-isothiocyanate, fluorescein 6-isothiocyanate or fluorescein thiosemicarbazide; the rhodamine is rhodamine 123, rhodamine B, rhodamine B isothiocyanate, rhodamine 6G, rhodamine 110 or ROX having the structure:



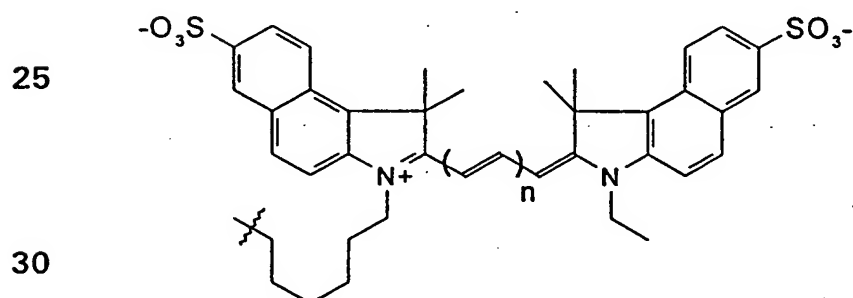
where each R is independently H or COOH; and the cyanine dye is Cy3, benzo-Cy3, dibenzo-Cy3, Cy5, benzo-Cy5, dibenzo-Cy5, Cy7, benzo-Cy7 or dibenzo-Cy7 having the structures:



$n = 1$ (Cy3), 2 (Cy5) or 3 (Cy7)



20 $n = 1$ (benzo-Cy3), 2 (benzo-Cy5) or 3 (benxo-Cy7)



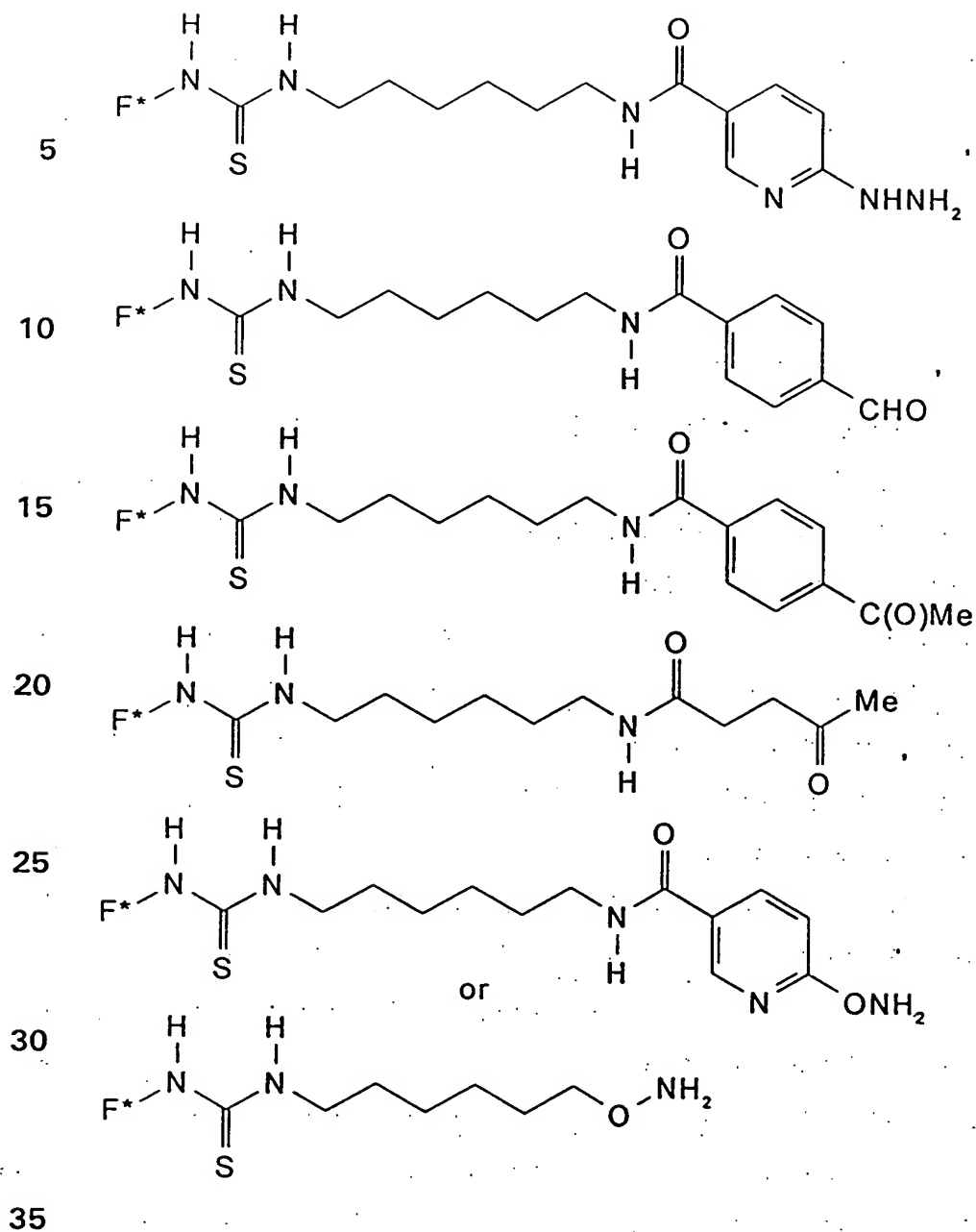
$n = 1$ (dibenzo-Cy3), 2 (dibenzo-Cy5) or 3 (dibenxo-Cy7)

35 See, e.g., FIGURE 12 and International Patent Application Publication No. WO 99/65993.

In more preferred embodiments, the modified indicator is chosen from

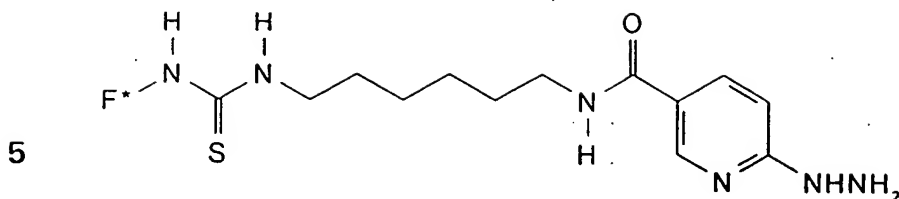
40

45



or a derivative thereof, where F* is selected as above.

More preferably, the modified indicators are those where F* is fluorescein, rhodamine, ROX, Cy3 or Cy5. Most preferably, the modified indicators have the formula:



where F* is fluorescein, rhodamine, ROX, Cy3 or Cy5.

D. Methods of Use of Modified Oligonucleotides

- 10 The hydrazino, oxyamino, and carbonyl modified oligonucleotides provided herein have application in numerous diagnostic methods. One application is use of the modified oligonucleotides in a DNA probe detection system. See, e.g., U.S. Patent No. 5,876,938. In such a system, a modified polynucleotide is prepared using random-primed DNA
- 15 labeling and a hydrazino, oxyamino or carbonyl containing monomer of formula (I) provided herein. It is to be understood that any of the monomers described herein could also be used, as well as alternative methods of polymer formation (e.g., nick translation, solid phase synthesis, and terminal transferase). Following preparation of the
- 20 hydrazino, oxyamino or carbonyl labeled probe, a blot hybridization can be carried out in which the labeled probe is applied to filter-bound DNA under conditions in which hybridization takes place between the probe and a target polynucleotide. The presence of a target polynucleotide can then be determined using, for example, enzyme-linked detection. Alkaline
- 25 phosphatase that is linked to a carbonyl, oxyamino or hydrazino moiety complementary to the hydrazino, oxyamino or carbonyl group of the modified oligonucleotide, will react with the hydrazino, oxyamino or carbonyl portion of the probe. Subsequent treatment with a substrate for alkaline phosphatase which forms a detectable product provides a means
- 30 by which the presence of the target nucleic acid or polynucleotide can be detected.

Alternatively, the modified oligonucleotides provided herein can be used for affinity purification of a target oligonucleotide. In one

embodiment, the probes are prepared as described above and added to a mixture of oligonucleotides containing a target oligonucleotide. Following hybridization, magnetic beads having an attached reactive moiety that is complementary to the hydrazino, oxyamino or carbonyl group of the probe are placed in the mixture and covalent bonding to the hydrazino, oxyamino or carbonyl labeled probe occurs via a hydrazone or oxime linkage. The beads (with attached probe and target) are drawn to a magnetic plate and the remaining materials are washed away.

In other embodiments, the modified oligonucleotides can be used in combination with other purification and labeling methods to provide unique methods of isolating and detecting picogram quantities of target oligonucleotides.

Accordingly, provided herein are methods for detecting the presence of a target nucleic acid in a sample, comprising;

(a) contacting a sample with a modified oligonucleotide which is substantially complementary to the target nucleic acid under conditions sufficient to hybridize the modified oligonucleotide to the target nucleic acid thereby forming a hybridized complex;

(b) contacting the hybridized complex with a complexing agent which comprises a detectable moiety or an indicator, and a carbonyl, oxyamino or hydrazino moiety that is complementary to the hydrazino, oxyamino or carbonyl group of the modified oligonucleotide; and

(c) detecting the presence of the detectable moiety or indicator, thereby detecting the presence of the target nucleic acid.

The hydrazino, oxyamino and carbonyl modified oligonucleotides or formula (II) provided herein are useful in a variety of other methods, including, but not limited to, diagnostic probe assays, DNA amplification by solid phase polymerase chain reactions (PCR), molecular computing (see, e.g., Adleman (1994) Science 266:1021-1024; Kari (1997)

Mathematical Intelligencer 19:9-22; Frutos et al. (1997) Nucleic Acids Res. 25:4748; Smith et al. (1998) J. Comp. Biol. 5:255; Liu et al. (1998)

J. Comp. Biol. 5:267; Frutos et al. (1998) J. Am. Chem. Soc. 120:10277; Wang et al. (1999) Biosystems 52:189-191; Liu et al. (1999) Biosystems 52:25-33; Liu et al. (2000) Nature 403:175-179; European Patent Application Publication No. EP 0 772 135; Reed et al. (June 2000) Scientific American:86-93), molecular addressing (Niemeyer et al. (1994) Nucl. Acids Res. 22(25):5530-5539), DNA sequencing by mass spectrometry (see, e.g., U.S. Patent Nos. 6,074,823 and 5,547,835) and in studying the molecular electronics of DNA (see, e.g., U.S. Patent Nos. 6,071,699, 6,066,448, 5,952,172 and 5,824,473). In general, the modified oligonucleotide is immobilized as described herein for use in the above methods.

The hydrazino modified oligonucleotides prepared using the reagents of formula (I) are also useful for a variety of additional purposes, including, but not limited to:

- 15 (i) direct attachment to surfaces such as carbonyl, i.e., aldehyde or ketone, modified glasses including controlled pore glass, plastics including polyethylene, polypropylene, polyamide and polyvinylidenedifluoride, celluloses including nitrocelluloses, latexes, metals including steel, gold, silver, aluminum, silicon and copper, silicas, agaroses including
- 20 Sepharose®, Sephadex® and polystyrenes, in the form of beads, capillaries, flat supports such as glass fiber filters, multiwell plates or membranes, wafers, combs, pins (e.g., arrays of pins suitable for combinatorial synthesis or analysis) or beads in pits of flat surfaces such as wafers (e.g., silicon wafers), with or without filter plates; resulting in
- 25 oligonucleotides attached via a hydrazone bond formed between the hydrazino oligonucleotide and the carbonyl moiety on the surface; and
- (ii) direct attachment to any biopolymer which has been modified to include a carbonyl moiety using, for example, periodate oxidation of carbohydrates or functionalization of amino moieties using carbonyl
- 30 modification reagents such as succinimidyl 4-formyl benzoate. See, e.g.,

Ghosh (European Patent Application Publication No. Ep 0 361 768), the disclosure of which is incorporated herein in its entirety.

In the cases wherein the final hydrazino product is required directly from standard aqueous ammonia deprotection and cleavage from a solid support, the hydrazino protecting group is a labile protecting group, including, but not limited to, FMOC, benzoyl or acetyl, that is removed during deprotection of the other oligonucleotide protecting groups. Thus direct addition of the crude reaction product to a solid surface appropriately modified with a moiety that reacts with the hydrazino function will lead to covalent modification of the surface, and non-hydrazino modified oligonucleotides and other impurities will be washed away.

E. Preparation of the Compounds

1. Monomers

In general, the monomers of formula (I) provided herein can be prepared by methods well known to those of skill in the art, or by modification of those methods using techniques known to those of skill in the art.

For example, monomers of formula (I) where X is a carbonyl group, particularly an aldehyde or ketone group, can be prepared starting from formyl- or keto-substituted carboxylic acids and carboxylic acid derivatives. Many such formyl- or keto-substituted carboxylic acids and carboxylic acid derivatives, including 4-formylbenzoic acid and 4'-carboxyacetophenone are commercially available (see, e.g., Aldrich Chemical Co., Milwaukee, Wisconsin). Reaction of the succinimidyl ester or other activated ester of these carboxylic acids, prepared, e.g., from the carboxylic acid, N-hydroxysuccinimide and dicyclohexylcarbodiimide, with, e.g., an amino or hydroxy substituted nucleoside triphosphate provides the desired carbonyl substituted nucleoside triphosphate (see, e.g., FIGURE 5).

In another example, the monomers of formula (I) where X is a hydrazino group or oxyamino group are prepared starting from hydrazino- or oxyamino-substituted carboxylic acids and carboxylic acid derivatives. Such hydrazino or oxyamino compounds are prepared by reaction of the
5 corresponding halocarboxylic acid with, e.g., hydrazine, hydroxylamine, or derivatives thereof. Semicarbazides and thiosemicarbazides are prepared by reaction of hydrazine with the corresponding isocyanates and isothiocyanates. Carbazides and thiocarbazides can be prepared by reaction of a hydrazine derivative with phosgene or thiophosgene, or an
10 equivalent thereof, and hydrazine. Hydrazides are prepared by reaction of an activated carboxylic acid with hydrazine. A carbonic acid dihydrazine is prepared by reaction of an isocyanate with carbonic acid dihydrazide. A hydrazine carboxylate can be prepared by reaction of an alcohol with phosgene and hydrazine.

15 The resulting hydrazino or oxyamino carboxylic acid is then protected. The protecting group is a salt or is any amine or hydrazine protecting group known to those of skill in the art.

Reaction of the succinimidyl ester or other activated ester of protected hydrazino or protected oxyamino carboxylic acids, prepared,
20 e.g., from the carboxylic acid, N-hydroxysuccinimide and dicyclohexylcarbodiimide, with, e.g., an amino or hydroxy substituted nucleoside triphosphate provides, after deprotection, the desired hydrazino or oxyamino substituted nucleoside triphosphate (see, e.g., FIGURES 7 and 8).

25 2. Modified oligonucleotides

The hydrazino, oxyamino or carbonyl-modified oligonucleotides provided herein can be synthesized in solid phase or in solution, using the above hydrazino, oxyamino or carbonyl-modified monomers and other nucleoside bases. In some embodiments, the hydrazino, oxyamino or
30 carbonyl-containing oligonucleotides are prepared using enzyme-based

methodology such as PCR, random prime labeling, tailing or nick translation.

The synthesis of modified polynucleotides containing a modified monomer described above can also be achieved by enzyme-based methods. Pyrimidine, purine and deazapurine, preferably cytosine, uracil, adenine, guanine and thymine, more preferably cytosine, nucleoside triphosphates containing a hydrazino, oxyamino or carbonyl moiety linked to the heterocyclic ring can be used as substrates for a wide variety of purified nucleic acid polymerases of both prokaryotic and eukaryotic origin. These include Taq DNA polymerase, DNA polymerase I of *E. coli*, bacteriophage T4 DNA polymerase, DNA polymerases alpha and beta from murine (A-9) and human (HeLa) cells, and the DNA polymerase of Herpes simplex virus. Nick-translation, random prime labeling, and terminal transferase tailing are also useful methods for the incorporation of a modified nucleic acid monomer into an polynucleotide. Nick-translation can be carried out as described in Rigby, et al. (1977) J. Mol. Biol. 113:237-251, incorporated herein by reference. Random prime labeling can be conducted utilizing a modification of the method of Feinberg, et al. (1988) Anal. Biochem. 132:6-13 in which a modified monomer is used in place of dTTP. Incorporation can be verified by capture of the probe (or modified oligonucleotide) on DHBHA-Sepharose. Tailing, or terminal transfer can be carried out using the method of Abhay, et al. (1988) Anal. Biochem. 169:376-382 in which a modified (boronic acid-containing) monomer is diluted into dTTP. As above, incorporation of the modified monomer can be verified by capture of the probe on DHBHA-Sepharose.

3. Modified indicators

The modified indicators provided herein can be prepared by the following methods, or by other methods known to those of skill in the art. In general, the commercially available (Amersham, BDL) fluorescent compound F* is derivatized as an amine, alcohol or thiol, preferably an

amine (see, e.g., FIGURE 6). Reaction of this F* derivative with an activated carboxylic acid compound, such as an acyl halide, a mixed anhydride, an aryl or heteroaryl ester, a naphthalimidyl ester, or a succinimidyl ester, preferably a succinimidyl ester, that is substituted with a protected hydrazino, protected oxyamino, or carbonyl group followed by deprotection provides the modified indicators of formula F*-M-X provided herein.

Alternatively, the fluorescent compound F* can be derivatized as an activated carboxylic acid compound, such as an acyl halide, a mixed anhydride, an aryl or heteroaryl ester, a naphthalimidyl ester, or a succinimidyl ester, preferably a naphthalimidyl ester. Reaction of this F* derivative with an amine, alcohol or carbonyl compound that possesses a protected or unprotected hydrazino, a protected or unprotected oxyamino, or carbonyl moiety affords, following deprotection, the modified indicators provided herein. Certain fluorescent compounds F* that are derivatized as activated carboxylic acids are known to those of skill in the art (see, e.g., International Patent Application Publication No. WO 99/65993, and U.S. Patent Nos. 4,337,063, 4,404,289, 4,405,711, 5,486,616, 5,569,587, 5,569,766 and 5,627,027).

In another alternative, an amino derivative of a fluorescent compound F* is converted to the corresponding isothiocyanate by reaction with, e.g., thiophosgene or thiocarbonyldiimidazole, and a tertiary amine, such as triethylamine or diisopropylethylamine. Reaction of this isothiocyanate derivative with an amine possessing a protected hydrazino, protected oxyamino or carbonyl group gives, following deprotection, the hydrazino, oxyamino and carbonyl modified indicators provided herein. Alternatively, reaction of the isothiocyanate derivative with an α,ω -diaminoalkane, such as 1,6-diaminohexane, followed by reaction with a protected hydrazino, protected oxyamino or carbonyl modified succinimidyl ester, as described herein, gives, after deprotection, the hydrazino, oxyamino or carbonyl modified indicators provided herein.

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLE 1

Synthesis of aldehyde triphosphate

5 5-Amino substituted cytidine triphosphate 5-1 was prepared from 2-deoxycytidine as described in U.S. Patent No. 5,242,796. As shown in FIGURE 5, triphosphate 5-1 was dissolved in 0.1 M phosphate, 0.15 M sodium chloride, pH 7.4 and diluted with DMF. A solution of succinimidyl 4-formylbenzoate (5-2; mmol, equiv, Pierce Chemicals, Rockford, IL) in
10 DMF was added to the triphosphate. The reaction mixture was incubated at room temperature and the progress of the reaction was followed by C-18 reverse phase HPLC (mobile phase 1: 50 mM triethylammonium acetate, pH 7.4, mobile phase 2: acetonitrile; gradient: 0% 2 to 20% 2 over 20 min, 20% 2 to 50% 2 over 5 min; flow: 2 mL/min). Further
15 reagent was added to drive the reaction to completion.

The product 5-3 was isolated by ion-exchange HPLC using DEAE-Sephadex with a gradient from water (500 mL) to 0.6 M LiCl (500 mL) as the eluting buffer. The compound was desalted by precipitation from acetone/MeOH. ³¹P NMR, ¹H NMR and mass spectral data confirmed the
20 structure of the product.

EXAMPLE 2

A. Synthesis of N-(6-fluoresceinyl)-N'-(6-aminohexyl)-thiourea

As shown in FIGURE 6, this compound was obtained by dissolving fluorescein isothiocyanate 6-1 (FITC)(1.03 g, 2.64 mmol) in 15 ml of
25 MeOH in a 50 ml oven dried flask. This formed an orange suspension that was allowed to stir for 5 min at room temperature after which 1,6-diaminohexane (1.23 g, 10.56 mmol) was added all at once and the solution instantly went homogeneous. This solution was allowed to stir at room temperature for 48 hr and was covered with aluminum foil after
30 which the reaction was completely evaporated in vacuo. The residue was dissolved in a solution of 1:1:3 EtOAc:MeOH:AcOH (4 ml). This

soln was loaded onto a column packed with 61 g of silica. The column had an inner radius of 2 cm. 500 ml of 1:1:2 EtOAc:MeOH:AcOH was pushed through the column by pressure and fractions were collected and analyzed by TLC. The product 6-2 had a R_f of 0.68 and this spot was collected and evaporated to dryness followed by 5 X 20 ml co-evaporation with MeOH. Yield: 1 g of an orange solid. 75% yield. Mass spec, NMR, and RP-HPLC all indicated pure material.

B. Synthesis of N-(6-fluoresceinyl)-N'-(6-(2-hydrazino-5-pyridylcarbonyl)aminohexyl)-thiourea hydrochloride

10 The product of Example 2.A 6-2 (174 mg, 0.344 mmol) was dissolved in 5 ml of dry DMF and treated with a solution of succinimidyl 2-BOC-hydrazino-5-pyridinecarboxylate (100 mg, 0.344mmol) in 2 ml of DMF was added by syringe all at once at room temperture. The reaction mixture was allowed to stir covered for 24 hr after which the reaction
15 was evaporated to dryness on rota vap. 4 M HCl in dioxane (40 ml) was added. A precipitate instantly formed. After 4 h, the precipitate was filtered and washed with 4 X 5 ml dioxane to yield 200 mg of 6-3 as an orange solid. 95% yield. ^1H NMR and mass spectrometry (calc 641.4; found 641.4) confirmed the structure.

20 **C. Synthesis of N-(6-fluoresceinyl)-N'-(6-(4-formylbenzoyl)amino-hexyl)-thiourea**

The product of Example 2.A 6-2 (100 mg, 0.198 mmol) was dissolved in a 1:2 solution of DMF:100 mM phosphate 1 mM NaCl buffer (pH 7.5)(10 ml). Succinimidyl 4-formylbenzoate (244 mg, 0.990 mmol) in
25 4 ml of dry DMF was added by syringe all at once. This solution was allowed to stir at room temperature for 48 hr, then was evaporated to dryness on rota-vap. The ensuing residue was treated with 40 ml of dioxane and allowed to sit overnight in the dark to induce precipitation. The precipitate was then filtered off and precipitated one more time with
30 EtO_2 to yield 110 g of pure yellow solid 6-4. 84% yield.

EXAMPLE 3

Incorporation of aldehyde triphosphate 5-3 in PCR reaction

<u>Oligonucleotide</u>	<u>Sequence</u>	<u>Source</u>
M13Fcomplement5'G	GGT CGT GAC TGG GAA AAC (SEQ ID NO:1)	Trilink
5 M13F	GTT TTC CCA GTC ACG AC (SEQ ID NO:2)	Genosys
M13R	AGC CCA TAA CAA TTT CAC ACA GGA (SEQ ID NO:3)	Genosys
TF3	AT GGA AGA ATC TCT CC (SEQ ID NO:4)	Genosys
Ferritin283	GGA ACA TGC TGA AAA CTG (SEQ ID NO:5)	Genosys
b-actin300	TCA CCA ACT GGG ACG ACA TG (SEQ ID NO:6)	Genosys
10	Oligonucleotide suppliers: Genosys, Inc (Woodlands, TX); Trilink Biotechnologies (San Diego, CA).	

The incorporation of aldehyde-dCTP into PCR generated double stranded DNA was performed using Stoffel Taq Polymerase (Perkin-Elmer) using buffer provided by the manufacturer and 3.5 mM MgCl₂.

- 15 Deoxynucleotides dGTP, dATP and TTP were present at 0.6 mM while primers M13 and M13 R (see sequences above) were used at 1 μ M. The DNA template was plasmid 80109 from the IMAGE Consortium collection of human cDNAs. PCR used different amounts of dCTP and aldehyde-dCTP at combined concentration of 0.6 mM. Individual reactions contained
- 20 0, 0.6, 0.18 and 0.6 mM aldehyde d-CTP. Syntheses of PCR products from each reaction was determined by radioincorporation of ³²P into Qiagen purified PCR product. The PAGE gel of the radiolabelled products demonstrates that reactions that included 10% and 30 % aldehyde-dCTP yielded full length product.

EXAMPLE 4

Fluorescent labeling of aldehyde incorporated PCR product

Determination of incorporation of aldehyde-dCTP was performed by incubation of hydrazino-fluorescein 6-3 with Qiagen purified PCR products (from Example 3) in 0.1 M acetate, pH 5.0 for 2 hours at room temperature. The labeling reaction was terminated by addition of one volume of 0.1 M Tris-HCl, pH 7.5. The sample was purified twice with a Qiagen PCR purification kit and eluted into a final volume of 25 μ L in EB buffer provided by Qiagen. Fluorescent signal was assayed by spotting 2 μ L of sample onto a microscope slide (Gold Seal, Fisher Scientific) and analysis using a ScanArray Fluorescent Scanner.

EXAMPLE 5

Incorporation of aldehyde-triphosphate by a variety of enzymes using a single-nucleotide extension assay

15	<u>Oligonucleotide</u>	<u>Sequence</u>	<u>Source</u>
	M13F complement 5'G	GGT CGT GAC TGG GAA AAC (SEQ ID NO:1)	Trilink
	M13F	GTT TTC CCA GTC ACG AC (SEQ ID NO:2)	Genosys
	M13R	AGC CCA TAA CAA TTT CAC ACA GGA (SEQ ID NO:3)	Genosys
	TF3	AT GGA AGA ATC TCT CC (SEQ ID NO:4)	Genosys
20	Ferritin283	GGA ACA TGC TGA AAA CTG (SEQ ID NO:5)	Genosys
	b-actin300	TCA CCA ACT GGG ACG ACA TG (SEQ ID NO:6)	Genosys

Oligonucleotide suppliers: Genosys, Inc (Woodlands, TX); Trilink 5 Biotechnologies (San Diego, CA)

³²P-labelling: Oligonucleotides were radiolabeled with ³²P using gamma ³²P-ATP 3000 Ci/mmol at 10 Ci/mL (ICN, Inc.). The labeling reaction was performed with 20 picomoles Oligonucleotide, 60 μ Ci gamma ³²P-ATP, 10 units T4 polynucleotide kinase (Gibco BRL) and 1 μ L 10X reaction buffer provided by the manufacturer. The reaction was incubated for 20 minutes at 37 °C. ³²P labeled oligonucleotides were

purified using a Qiagen nucleotide purification kit and samples are eluted into 50 μ L dH₂O and stored -20 °C.

Extension Assay

Oligonucleotides M12F complement 5'G and M13F were combined at a final concentration of 4 μ M along with 8.5X10⁶ cpm 32 labeled M13F in a final volume of 16 μ L. The mixture was incubated at 80 °C for 3 minutes and allowed to cool to room temperature. Single nucleotide extension assays were performed using a variety of enzymes, *i.e.*, T4 DNA polymerase, Klenow, E. coli DNA polymerase I, Stoffel Taq polymerase, Superscript reverse transcriptase, MMLV reverse transcriptase, and AMV reverse transcriptase. Reaction conditions used buffers supplied by the manufacturers. Reactions also contained either dCTP or aldehyde-dCTP. Extension reactions were stopped by freezing at -20 °C. Analysis of extension reactions was performed by polyacrylamide gel electrophoresis on a 12.5% polyacrylamide gel containing 8M urea.

EXAMPLE 6

Synthesis of hydrazine triphosphate

5-Amino substituted cytidine triphosphate 7-1 was prepared from 2-deoxycytidine as described in U.S. Patent No. 5,242,796. As shown in FIGURE 7, triphosphate 7-1 was dissolved in 0.1 M phosphate, 0.15 M sodium chloride, pH 7.4 and diluted with DMF. A solution of succinimidy 6-hydrazinonicotinate acetone hydrazone (7-2; 5 equiv, Pierce Chemicals, Rockford, IL) in DMF was added to the triphosphate. The reaction mixture was incubated at room temperature and the progress of the reaction was followed by C-18 reverse phase HPLC (mobile phase 1: 50 mM triethylammonium acetate, pH 7.4, mobile phase 2: acetonitrile; gradient as in Example 1). Further reagent was added to drive the reaction to completion. The product 7-3 is isolated by ion-exchange HPLC using conditions as described in Example 1.

EXAMPLE 7

Synthesis of aminooxy triphosphate

As shown in FIGURE 8, a solution of N-hydroxyphthalimide (1 equiv; 8-1; Aldrich Chemical Co., Milwaukee, WI) is dissolved in DMF and treated with triethylamine (1 equiv) and 5-bromopentanoic acid (1 equiv.; Aldrich Chemical Co.). The reaction mixture is heated at 50 °C and the progress of the reaction followed by silica gel thin layer chromatography until complete. The solvent is removed under reduced pressure and the residue is partitioned between ethyl acetate and 5% citric acid. The organic phase is washed with brine, dried (magnesium sulfate, filtered and concentrated to the crude product. The product 8-2 is purified by silica gel chromatography.

To a solution of acid 8-2 (1 equiv) in THF is added N-hydroxysuccinimide (1 equiv; Aldrich Chemical Co.) and a solution of dicyclohexylcarbodiimide (1 equiv; Aldrich Chemical Co.) The reaction mixture is stirred at room temperature for 4 hours and the precipitate that forms (dicyclohexylurea) is removed by filtration. The filtrate is concentrated to dryness and the residue is recrystallized to yield desired succinimidyl ester 8-3.

5-Amino substituted cytidine triphosphate 8-4 (prepared from 2-deoxycytidine according to the procedure of U.S. Patent No. 5,242,796) is dissolved in 0.1 M phosphate, 0.15 M sodium chloride, pH 7.4 to prepare a 10 mg/mL solution and diluted with an equal volume of DMF. A solution of 8-3 (1 equiv) in DMF is added to the triphosphate solution. The reaction mixture is incubated at room temperature and the progress of the reaction is followed by C-18 reverse phase HPLC (mobile phase 1: 50 mM triethylammonium acetate, pH 7.4, mobile phase 2: acetonitrile; gradient as described in Example 1). Further 8-4 is added to drive the reaction to completion. The product 8-5 is isolated by ion-exchange HPLC. ³¹P NMR, ¹H NMR and mass spectral data confirmed the structure of the product.

Phthalimide 8-5 is converted into oxyamino 8-6 by addition of 8-5 to a solution of hydrazinium acetate. The reaction mixture is incubated at room temperature. The purified product is isolated by ion exchange chromatography.

5

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.